

2007-10-15 15:07

809 CAGGATCTCAGGACCCCTGGTGCCTCTGTGGAGAGAGAACCGGCCACCGAGCTGCT 1888
 1827 GTTAACACCTTAATTCGAGAAAGTGTGTCTGTATCTTGTCTGTATCTACACAGG 1586
 889 GTTAACAGCTTAATTCGAGAAAGTGTGTCTGTATCTTGTCTGTATCTACACAGG 948
 1587 CCGACGGCTTGGAGAGCTTCACAGGTGTGTGAGACCGCCAGCCATGTGGAGATCCTG 1646
 949 CCGACGGCTTGGAGAGCTTCACAGGTGTGTGAGACCGCCAGCCATGTGGAGATCCTG 1508
 1647 GGTCTCTCTAGGAGAGAAAGATATCTTTACAGATCTTTCACATATCAGAGAGAG 1706
 1609 GGTCTCTCTAGGAGAGAAAGATATCTTTACAGATCTTTCACATATCAGAGAGAG 1649
 1727 GGTGACCAAGCTTTCATTTAGTGAAGTGAAGAGGCTTCTGTACATATCTACTCGTC 1766
 1549 GCGGACCAATCTTCAATTAAGTGAAGATACAGAGCTTCTGTACATATCTACTCGTC 1128
 1767 CCGCTGTGTCTGTGGTGTGTGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1826
 1129 CCGCTGTGTCTGTGGTGTGTGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1168
 1817 TTGAGACAGCTCCAGAGACACATATTAAGATTAAGAGGCTTGTGTGTGTGTGTGTGT 1886
 1249 TTGAGACAGCTCCAGAGACACACATTAAGATTAAGAGGCTTGTGTGTGTGTGTGTGT 1248
 1887 CAACCCAGCCGGGGCCCGCGGTCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1946
 1249 CAACCCAGCCGGGGCCCGCGGTCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1308
 1917 TTGGGACAGATGT 2006
 1309 TTGGGACAGATGT 1368
 2002 CAGGACAG 2066
 1349 CAGGACAG 1426
 2002 ATCAACTGTAG 2078
 1449 ATCAACTGTAGAGCTACTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1488
 2078 2278
 1449 ATCAACTGTAGAGCTACTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1548
 2078 2278
 1549 GGT 2111
 2112 TTGGAGCTGTGAAG 2171
 1609 TTGGAGCTGTGAAG 1668
 2172 TGGCGGACATCAGAGATCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2231
 1449 TGGCGGACATCAGAGATCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1728
 2232 TGGCGGACATCAGAGATCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2291
 1729 TGGCGGACATCAGAGATCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1788
 2292 GAGTGTATCCAG 2351
 1789 GAGTGTATCCAG 1848
 2352 AAGATGAGAGACATGGT 2411
 1849 AAGATGAGAGACATGGT 1908
 2412 CACTGTATGGCGCACCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2471

1909 CACTGTATGTAGCGCACCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1968
 2472 SGGTACATCTGT 2528
 1969 SGGTACATCTGT 2528
 2529 GAATATCTGT 2588
 2029 GAATATCTGT 2088
 2589 AATTCCTGT 2648
 2089 AATTCCTGT 2148
 2649 AAATCTGT 2708
 2149 AAATCTGT 2164
 2709 TTGGTCAAGATGT 2768
 2165 2165
 2769 GGATCTGT 2828
 2165 2165
 2829 CTAATATGT 2888
 2165 2165
 2889 AGGTATGT 2948
 2165 2165
 2949 GCACTGT 3008
 2165 2165
 3009 CAAATATGT 3068
 2240 TAAATATGT 2299
 3069 GGT 3128
 2300 GGT 2459
 3129 GGT 3188
 2360 GGT 2419
 3189 GGT 3248
 2420 GGT 2479
 3249 ACCAGT 3308
 2480 ACCAGT 2539
 3309 GGT 3368
 2540 GGT 2599
 3369 TGAATGT 3428
 2600 TGAATGT 2669
 3429 GGT 3488
 2660 GGT 2728
 3489 GGT 3548
 2660 GGT 2728

255 ----- GGGTGAAGAGTGGCCCATCTCCAGCTCAGCTGGCAGG 793

2973 ACCTCTGCGAGCTCTATAGCCGATGAGATTTGAAATGATGATCTCAGTGGCAGG 3032

2994 ACCTCTGCGAGCTCTATAGCCGATGAGATTTGAAATGATGATCTCAGTGGCAGG 853

3035 GCGTGGATCCCAAGGATGAGTCTCTTGGAGGCTCGGCGATGCGCAATGAGG 3092

854 GCGTGGATCCCAAGGATGAGTCTCTTGGAGGCTCGGCGATGCGCAATGAGG 913

3093 TGGAGATGATTCAGTGGAGAGTCTGATGAGATGAGGATGATGAGATGAGCTT 3152

914 TGGAGATGATTCAGTGGAGAGTCTGATGAGATGAGGATGATGAGATGAGCTT 973

3153 GCGTGGATCCCAAGGATGAGTCTCTTGGAGGCTCGGCGATGCGCAATGAGG 3212

974 GCGTGGATCCCAAGGATGAGTCTCTTGGAGGCTCGGCGATGCGCAATGAGG 1033

3213 ATTGGGCTGAGTCTTACTATCGAGGAGTCTGATGAGATGAGGATGAGCTT 3272

1034 ATTGGGCTGAGTCTTACTATCGAGGAGTCTGATGAGATGAGGATGAGCTT 1093

3273 TGTGATGAGATCTGCGGCTGATCTGATGAGATGAGGATGAGGATGAGCTT 3332

1094 TGTGATGAGATCTGCGGCTGATCTGATGAGATGAGGATGAGGATGAGCTT 1153

3333 GGTGAGGAGGAGCTGAGAGAGTCTGATGAGATGAGGATGAGGATGAGCTT 3392

1154 GGTGAGGAGGAGCTGAGAGAGTCTGATGAGATGAGGATGAGGATGAGCTT 1213

3393 TGTGATGAGATCTGCGGCTGATCTGATGAGATGAGGATGAGGATGAGCTT 3452

1214 TGTGATGAGATCTGCGGCTGATCTGATGAGATGAGGATGAGGATGAGCTT 1273

3453 TC 3454

1274 GC 1275

RESULT 5

AA568757 standard: CENA; 526 BP.

XX AA568757:

OT 11-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein: 8461.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical; imaging; diagnostic; genetic disorders; ss.

XX H577 suppliers.

XX WOL200175067-32.

XX 11-OCT-2001.

XX 10-MAR-2001; 2001WC-US88631.

XX 31-MAR-2003; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX CHYSE-1 HYSEQ INC.

XX Datanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; AB044570.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess

biodiversity

Claim 1: SEQ ID No 4561: 103pp; English.

The invention relates to isolated polynucleotide (i) and

polypeptide (ii) sequences, (ii) is useful as hybridisation probes,

polymerase chain reaction (PCR) primers, oligomers, and for chromosome

and gene mapping, and in recombinant production of (ii). The

polynucleotides are also used in diagnostics as expressed sequence tags

for identifying expressed genes. (i) is useful in gene therapy techniques

to restore normal activity of (ii) or to treat disease states involving

(ii). (ii) is useful for generating antibodies against it, detecting or

quantifying a polypeptide in tissue, as molecular weight markers and as

a food supplement. (ii) and its binding partners are useful in medical

imaging of sites expressing (ii). (i) and (ii) are useful for treating

disorders involving aberrant protein expression or biological activity.

The polypeptide and polynucleotide sequences have applications in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits to assess biodiversity

and to produce other types of data and products dependent on RNA and

amino acid sequences. AA564197 AA594564 represent novel human

diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed

specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pat_sequences.

XX

SQ Sequence 506 BP; 108 A; 149 C; 148 G; 101 T; 0 other:

Query Match 12.0%; Score 420.9; DB 23; Length 506;

Best Local Similarity 99.4%; Prod. No. 1.4e-86;

Matches 503; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2085 TTCTTGGCAATACACAGGCTCTCTCTTGGACTCTGACAGGAGGAGACAGACAC 2144

DB 1 TTCTTGGCAATACACAGGCTCTCTCTTGGACTCTGACAGGAGGAGACAGACAC 63

QY 2145 CTGGAGAGAGTTCTCTTGGAGGTTGGTCCATCATAGATGGACCTGTTGGAGTGG 2184

DB 61 CTGGAGAGAGTTCTCTTGGAGGTTGGTCCATCATAGATGGACCTGTTGGAGTGG 120

QY 2225 ATGCAAAACCAAGATTCAGAGGCTGCTACCTGACAGGAGCTCTTGGAGTGGTTC 2164

DB 121 ATGCAAAACCAAGATTCAGAGGCTGCTACCTGACAGGAGCTCTTGGAGTGGTTC 180

QY 2225 AGCTCTCTTGAAGATTCAGAGGAGGAGGATTTATTCAGAGGAGGCTGAGGAGTGGAG 2424

DB 181 AGCTCTCTTGAAGATTCAGAGGAGGAGGATTTATTCAGAGGAGGCTGAGGAGTGGAG 246

QY 2325 GTGATCTGTCTCAGCAATATGCTCTCAGATGAGGACATGGCTCCCTGCTGTGTCTCT 2183

DB 241 GTGATCTGTCTCAGCAATATGCTCTCAGATGAGGACATGGCTCCCTGCTGTGTCTCT 300

QY 2384 GAAGTGTCTGAGGAGGCTGAGGCTGCTGCACTTGTATGGGCGCCACCTACAGCGGAGCG 2443

DB 301 GAAGTGTCTGAGGAGGCTGAGGCTGCTGCACTTGTATGGGCGCCACCTACAGCGGAGCG 303

QY 2444 GGAGATCTGAGGAGGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2503

DB 361 GGAGATCTGAGGAGGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420

QY 2504 GACCTCTCTGCTGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2563

DB 421 GACCTCTCTGCTGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

QY 2564 CTTGATAGAGTGTCTCTGTACCCAA 2589

DB 481 CTTGATAGAGTGTCTCTGTACCCAA 506

RESULT 6

AACT6566

ID AAC76566 standard: CENA; 409 BP.

PR 26-JUL-2003; 2000US-0220963.
PR 26-JUL-2003; 2000US-0220964.
PR 14-AUG-2003; 2000US-0224518.
PR 14-AUG-2003; 2000US-0224519.
PR 14-AUG-2003; 2000US-0225211.
PR 14-AUG-2003; 2000US-0225214.
PR 14-AUG-2003; 2000US-0225266.
PR 14-AUG-2003; 2000US-0225267.
PR 14-AUG-2003; 2000US-0225268.
PR 14-AUG-2003; 2000US-0225270.
PR 14-AUG-2003; 2000US-0225447.
PR 14-AUG-2003; 2000US-0225757.
PR 14-AUG-2003; 2000US-0225758.
PR 14-AUG-2003; 2000US-0225759.
PR 18-AUG-2003; 2000US-0226279.
PR 22-AUG-2003; 2000US-0226681.
PR 22-AUG-2003; 2000US-0226685.
PR 22-AUG-2003; 2000US-0227182.
PR 23-AUG-2003; 2000US-0227209.
PR 30-AUG-2003; 2000US-0228924.
PR 01-SEP-2003; 2000US-0229287.
PR 01-SEP-2003; 2000US-0229343.
PR 01-SEP-2003; 2000US-0229344.
PR 01-SEP-2003; 2000US-0229345.
PR 03-SEP-2003; 2000US-0229509.
PR 06-SEP-2003; 2000US-0229513.
PR 06-SEP-2003; 2000US-0230437.
PR 06-SEP-2003; 2000US-0230438.
PR 08-SEP-2003; 2000US-0231242.
PR 08-SEP-2003; 2000US-0231243.
PR 08-SEP-2003; 2000US-0231244.
PR 08-SEP-2003; 2000US-0231413.
PR 08-SEP-2003; 2000US-0231414.
PR 08-SEP-2003; 2000US-0232080.
PR 12-SEP-2003; 2000US-0232081.
PR 14-SEP-2003; 2000US-0232397.
PR 14-SEP-2003; 2000US-0232398.
PR 14-SEP-2003; 2000US-0232399.
PR 14-SEP-2003; 2000US-0232400.
PR 14-SEP-2003; 2000US-0232401.
PR 14-SEP-2003; 2000US-0233063.
PR 14-SEP-2003; 2000US-0233064.
PR 14-SEP-2003; 2000US-0233065.
PR 21-SEP-2003; 2000US-0234223.
PR 21-SEP-2003; 2000US-0234274.
PR 25-SEP-2003; 2000US-0234997.
PR 25-SEP-2003; 2000US-0234998.
PR 26-SEP-2003; 2000US-0235484.
PR 27-SEP-2003; 2000US-0235834.
PR 27-SEP-2003; 2000US-0235836.
PR 29-SEP-2003; 2000US-0236327.
PR 29-SEP-2003; 2000US-0236367.
PR 29-SEP-2003; 2000US-0236368.
PR 29-SEP-2003; 2000US-0236369.
PR 29-SEP-2003; 2000US-0236370.
PR 02-OCT-2003; 2000US-0236802.
PR 02-OCT-2003; 2000US-0237037.
PR 02-OCT-2003; 2000US-0237038.
PR 02-OCT-2003; 2000US-0237039.
PR 02-OCT-2003; 2000US-0237640.
PR 13-OCT-2003; 2000US-0239935.
PR 13-OCT-2003; 2000US-0239937.
PR 20-OCT-2003; 2000US-0240960.
PR 20-OCT-2003; 2000US-0241221.
PR 20-OCT-2003; 2000US-0241785.
PR 20-OCT-2003; 2000US-0241786.
PR 20-OCT-2003; 2000US-0241787.
PR 20-OCT-2003; 2000US-0241808.
PR 20-OCT-2003; 2000US-0241809.
PR 20-OCT-2003; 2000US-0241826.
PR 01-NOV-2003; 2000US-0244617.
PR 08-NOV-2003; 2000US-0246474.

PR 08-NOV-2003; 2000US-0246475.
PR 08-NOV-2003; 2000US-0246476.
PR 08-NOV-2003; 2000US-0246477.
PR 08-NOV-2003; 2000US-0246478.
PR 08-NOV-2003; 2000US-0246523.
PR 08-NOV-2003; 2000US-0246524.
PR 08-NOV-2003; 2000US-0246525.
PR 08-NOV-2003; 2000US-0246526.
PR 08-NOV-2003; 2000US-0246527.
PR 08-NOV-2003; 2000US-0246528.
PR 08-NOV-2003; 2000US-0246529.
PR 08-NOV-2003; 2000US-0246530.
PR 08-NOV-2003; 2000US-0246611.
PR 08-NOV-2003; 2000US-0246613.
PR 17-NOV-2003; 2000US-0249207.
PR 17-NOV-2003; 2000US-0249208.
PR 17-NOV-2003; 2000US-0249209.
PR 17-NOV-2003; 2000US-0249210.
PR 17-NOV-2003; 2000US-0249211.
PR 17-NOV-2003; 2000US-0249212.
PR 17-NOV-2003; 2000US-0249213.
PR 17-NOV-2003; 2000US-0249214.
PR 17-NOV-2003; 2000US-0249215.
PR 17-NOV-2003; 2000US-0249216.
PR 17-NOV-2003; 2000US-0249217.
PR 17-NOV-2003; 2000US-0249218.
PR 17-NOV-2003; 2000US-0249244.
PR 17-NOV-2003; 2000US-0249245.
PR 17-NOV-2003; 2000US-0249246.
PR 17-NOV-2003; 2000US-0249255.
PR 17-NOV-2003; 2000US-0249297.
PR 17-NOV-2003; 2000US-0249299.
PR 17-NOV-2003; 2000US-0249300.
PR 01-DEC-2003; 2000US-0251160.
PR 05-DEC-2003; 2000US-0251030.
PR 05-DEC-2003; 2000US-0251988.
PR 05-DEC-2003; 2000US-0256715.
PR 08-DEC-2003; 2000US-0251479.
PR 08-DEC-2003; 2000US-0251856.
PR 08-DEC-2003; 2000US-0251858.
PR 08-DEC-2003; 2000US-0251859.
PR 08-DEC-2003; 2000US-0251949.
PR 11-DEC-2003; 2000US-0251990.
PR 05-JAN-2004; 2001US-0254097.
PR 05-JAN-2004; 2001US-0259678.

!HUYA-! HUMAN_GENOME_SCI :NC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and
metastasis -

Disclosure: SEQ ID NO 25807; 307pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (!)
amino acid sequences given in AAK62170 to AAK91921. (!) have catalytic
activity, and can be used in gene therapy and vaccine production. (!)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (!) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (!) by expressing inactive proteins or to
supplement the patient's own production of (!). Additionally, (!)
polynucleotides may be used to produce the secreted (!), by inserting
the nucleic acids into a host cell and culturing the cell to express the
protein. (!) proteins and polynucleotides may be used to prevent,
diagnose and treat immune/haematopoietic-related diseases, especially

CC cancers and cancer metastases of hematopoietic-derived cells. AAR64703
 to AAR67694 represent human immune/zev-4 epitope to anti-genomic
 CC sequences from the present invention. AAK54942 to AAK54952 and AAK52169
 CC represent sequences used in the experimental of the present invention.

XX Sequence 19199 BP: 4465 A; 4246 G; 4865 G; 5023 T; 0 other;
 Query Match: 8.41; Score 4157; E-Val 22; Length 19199;
 Best Local Similarity: 55.04; P-Val 1.7e-71;
 Matches 1123; Conservative 0; Mismatches 860; Indels 222; Gaps 17;
 2675 TTGAATTAATATATTGGCGAGTATATGCTTACAGCTTAAATTCACACACATTTTGGGA 2754
 11118 TTGAAGAATATTACAGCTGGGACGATGAGCTTACGCTTAAATTCACAGCTTTTGGGA 11167
 2755 GCCCAGATGGGAGATCAATTGATGAGAGTTCAGACAGATCTGGCGAATAGGTS 2814
 11168 GCGTGGCGGCGAGTCACTTGATGAGTTCAGAGTTCGAGATCAGCTTGACAGCTGGAG 11227
 2815 AACCCCATCTCTACTAAATAT-ATTAAAT-AGG-CAAG-ATGATGAGACAGCTGTATA 2873
 11228 AACCCCATCTCTACTAAATAT-ATTAAAT-AGG-CAAG-ATGATGAGACAGCTGTATA 11287
 2874 GCCCATCTCTACTAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2933
 11288 TCCCATCTCTACTAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 11347
 2934 GATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2992
 11348 GATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 11407
 2993 GCGCATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3052
 11408 AAAAAACACAAAAATTTACAAAGATTTTAAAGAAATCAGAGTACGATGACTACTGTT 11467
 3053 ATGCTGTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3112
 11468 GATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 11527
 3113 AAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3172
 11528 AATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 11581
 3173 ATCTGCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3232
 11582 GAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 11641
 3233 TGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3287
 11642 TATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 11701
 3288 GTCGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3347
 11702 ACCGCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 11761
 3348 TGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3407
 11763 TCTTACT 3466
 3408 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3467
 11802 GT 11861
 3468 CGGTTTCTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3527
 11862 AGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 11921
 3494 SCAGCGCTCTACCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3553
 11942 TCT 12001
 3554 TACATCAGCCCT 3613

DB 12002 TCCCGGCTTCCGCCATTTCTGCTCAGCGCTCCCAAGTAGCTGGGATTACAGGCACAT 12061
 QY 3614 GCCACCAAGCCAGCTAAATTTTCTGATTTTATAGTAGAGACAGGTTTTCATCATGTTAG- 3672
 DB 12062 GCACACAGCCGCTAGTATTTTATATTTTATAGTAGAGTGGGTTTTCATGTTTANA 12121
 QY 3673 CAGGATGTTCTGATTTACTAGCTCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3732
 DB 12122 CAGATGTTTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12181
 QY 3733 GATTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCT 3789
 DB 12182 GATTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCT 12241
 QY 3770 GCTTTTAA 3778
 DB 12242 GCATTAAATTCATTTCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 12301
 QY 3779AAGGATTTTCTGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3831
 DB 12302 GACAAATTAAGAGAGACCTGAGGCGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12361
 QY 3832 GCGAGGCTGAGGCTGGAGGAGTCCAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3891
 DB 12362 GCGAGGCTGAGGCTGGAGGAGTCCAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 12421
 QY 3892 GTAGAAATACGCTATTTCTTACTTAAATATCAAAATAGCCAGGATGCTGCTGCTGCTGCT 3951
 DB 12422 GGAATATCTGCTATCTATCAAAATCAAAATAGCCAGGATGCTGCTGCTGCTGCTGCT 12479
 QY 3952 ATAAATCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4011
 DB 12480 ATAAATCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12539
 QY 4012 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4071
 DB 12540 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12599
 QY 4072 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4131
 DB 12600 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12657
 QY 4132 ATTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4191
 DB 12658AGTAACTGAAATGAGG 12675
 QY 4192 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4251
 DB 12676 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12735
 QY 4252 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4311
 DB 12736 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12795
 QY 4312 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4371
 DB 12796 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12855
 QY 4372 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4431
 DB 12856 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12915
 QY 4432 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4491
 DB 12916 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12975
 QY 4492 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4551
 DB 12976 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13035
 QY 4552 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4611
 DB 13036 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13095

CY	4884AAACAAAAACAAAAATCAACACGGCGCTGGCGCGTG	4828	PR	14-AUG-2000	2000US-0205266
EB	13236	AATTCATAGGTTTCAGTGAAGAAAGAAAGAGATCTCAGAGTTTGGGCAAGTG	13155	PR	14-AUG-2000	2000US-0225267
QY	4829	GCTGATGCTGTAAACCGACACTTTTGGAGAGCTGAGAGCTGGTGGTGGTAC	4866	PR	14-AUG-2000	2000US-0225268
DE	13156	TCTCATCTTTGTAATCCAGCAATTTAGGAGGCTAGGCTGAGAGGATGCTTGAAGCCAG	13215	PR	14-AUG-2000	2000US-0225270
QY	4832	GAGATGGAGACATCTCTGGGAC	4744	PR	14-AUG-2000	2000US-0225477
DE	13216	AAGTTTGACATCAGCTTAGGCAATCGGTAAGA	13235	PR	14-AUG-2000	2000US-0225478
QY	4845	TGACTTGGAGCTGTGGTGTACATGTATGAGTACTACAGAGTTGATGCAAG	4804	PR	14-AUG-2000	2000US-0225479
DE	13236	TGACTTGGAGCTGTGGTGTACATGTATGAGTACTACAGAGTTGATGCAAG	13335	PR	14-AUG-2000	2000US-0225480
QY	4865	AATTCATTTGACACGAGGAGGAGGTTTCACTTACGCAATATGAGTACTTCCA	4864	PR	14-AUG-2000	2000US-0225481
DE	13336	GATCATTGACCTGGGAGGTCAGTCTGTAAATTAATGATGATTCCTCCTACTCTA	13395	PR	14-AUG-2000	2000US-0225482
QY	4865	GCTTGGGACAGAGTGGATTCATTTAAAAAAGAAAAAAGAAAAAAGAAAAA	4924	PR	14-AUG-2000	2000US-0225483
DE	13396	GCTTGGGACAGAGTGGATTCATTTAAAAAAGAAAAAAGAAAAAAGAAAAA	14555	PR	14-AUG-2000	2000US-0225484
QY	4945	AAAAA 4929		PR	14-AUG-2000	2000US-0225485
DE	13456	AAAAA 13460		PR	14-AUG-2000	2000US-0225486
RESULT 4						
AAAT-42872	AAAT-42872 standard; DNA: 26277 BP.					
XX	AAAT-42872	AAAT-42872				
XX	AAAT-42872	AAAT-42872				
DT	06 NOV-2001	first entry:				
XX	Human immune/haematopoietic antigen gene/sequence SEQ ID NO:26242.	Human immune/haematopoietic antigen gene/sequence SEQ ID NO:26242.				
XX	Human immune/haematopoietic antigen gene/sequence SEQ ID NO:26242.	Human immune/haematopoietic antigen gene/sequence SEQ ID NO:26242.				
XX	Cytostatic gene therapy; vaccine. Not for clinical use.	Cytostatic gene therapy; vaccine. Not for clinical use.				
CS	Human sapiens.	Human sapiens.				
XX	W000147182-A2.	W000147182-A2.				
XX	09-AUG-2001.	09-AUG-2001.				
XX	17 JAN-2001: 2001WC-US01354.	17 JAN-2001: 2001WC-US01354.				
XX	31 JAN-2000: 2000US-0179065.	31 JAN-2000: 2000US-0179065.				
XX	04 FEB-2000: 2000US-0180628.	04 FEB-2000: 2000US-0180628.				
XX	24 FEB-2000: 2000US-0184664.	24 FEB-2000: 2000US-0184664.				
XX	02 MAR-2000: 2000US-0186353.	02 MAR-2000: 2000US-0186353.				
XX	14 MAR-2000: 2000US-0188874.	14 MAR-2000: 2000US-0188874.				
XX	17 MAR-2000: 2000US-0190376.	17 MAR-2000: 2000US-0190376.				
XX	18 APR-2000: 2000US-0198123.	18 APR-2000: 2000US-0198123.				
XX	19 MAY-2000: 2000US-0205515.	19 MAY-2000: 2000US-0205515.				
XX	07 JUN-2000: 2000US-0209467.	07 JUN-2000: 2000US-0209467.				
XX	28 JUN-2000: 2000US-0214866.	28 JUN-2000: 2000US-0214866.				
XX	30 JUN-2000: 2000US-0215325.	30 JUN-2000: 2000US-0215325.				
XX	07 JUL-2000: 2000US-0216647.	07 JUL-2000: 2000US-0216647.				
XX	07 JUL-2000: 2000US-0216880.	07 JUL-2000: 2000US-0216880.				
XX	11 JUL-2000: 2000US-0217497.	11 JUL-2000: 2000US-0217497.				
XX	11 JUL-2000: 2000US-0217496.	11 JUL-2000: 2000US-0217496.				
XX	14 JUL-2000: 2000US-0218290.	14 JUL-2000: 2000US-0218290.				
XX	26 JUL-2000: 2000US-0220963.	26 JUL-2000: 2000US-0220963.				
XX	26 JUL-2000: 2000US-0220964.	26 JUL-2000: 2000US-0220964.				
XX	14 AUG-2000: 2000US-0224518.	14 AUG-2000: 2000US-0224518.				
XX	14 AUG-2000: 2000US-0224519.	14 AUG-2000: 2000US-0224519.				
XX	14 AUG-2000: 2000US-0225213.	14 AUG-2000: 2000US-0225213.				
XX	14 AUG-2000: 2000US-0225214.	14 AUG-2000: 2000US-0225214.				

FR 08-NOV-2000; 2000US-0246529;
 PR 09-NOV-2000; 2000US-0246526;
 PR 09-NOV-2000; 2000US-0246527;
 PR 09-NOV-2000; 2000US-0246528;
 PR 09-NOV-2000; 2000US-0246532;
 PR 08-NOV-2000; 2000US-0246539;
 PR 08-NOV-2000; 2000US-0246610;
 PR 08-NOV-2000; 2000US-0246611;
 PR 09-NOV-2000; 2000US-0246613;
 PR 17-NOV-2000; 2000US-0249207;
 PR 17-NOV-2000; 2000US-0249208;
 PR 17-NOV-2000; 2000US-0249209;
 PR 17-NOV-2000; 2000US-0249210;
 PR 17-NOV-2000; 2000US-0249211;
 PR 17-NOV-2000; 2000US-0249212;
 PR 17-NOV-2000; 2000US-0249213;
 PR 17-NOV-2000; 2000US-0249214;
 PR 17-NOV-2000; 2000US-0249215;
 PR 17-NOV-2000; 2000US-0249216;
 PR 17-NOV-2000; 2000US-0249217;
 PR 17-NOV-2000; 2000US-0249218;
 PR 17-NOV-2000; 2000US-0249244;
 PR 17-NOV-2000; 2000US-0249245;
 PR 17-NOV-2000; 2000US-0249246;
 PR 17-NOV-2000; 2000US-0249265;
 PR 17-NOV-2000; 2000US-0249299;
 PR 17-NOV-2000; 2000US-0249299;
 PR 01-DEC-2000; 2000US-0249300;
 PR 01-DEC-2000; 2000US-0250391;
 PR 03-DEC-2000; 2000US-0251012;
 PR 03-DEC-2000; 2000US-0251988;
 PR 04-DEC-2000; 2000US-0256719;
 PR 06-DEC-2000; 2000US-0251479;
 PR 08-DEC-2000; 2000US-0251856;
 PR 08-DEC-2000; 2000US-0251868;
 PR 08-DEC-2000; 2000US-0251869;
 PR 08-DEC-2000; 2000US-0251989;
 PR 09-DEC-2000; 2000US-0251990;
 PR 11-DEC-2000; 2000US-0254097;
 PR 05-JAN-2001; 2001US-0259678;
 XX

PA (HUMA) HUMAN GENOME SCI INC.

PI Robert CA, Karasik SC, Ruben SM;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 9.3%; Score 410.4; DB 22; Length 26277;
 Best Local Similarity 61.3%; Field No. 14e 70;
 Matches 885; Conservative 0; Mismatches 511; Indels 38; Gaps 13;
 QY 3530 TCTCTGATCCCAACCTTAGCTATTATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3599
 DB TCTCTGATCCCAACCTTAGCTATTATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3599
 QY 3590 AATGATGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3649
 DB AATGATGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3649
 QY 2193 AATGATGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 2195
 DB AATGATGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 2195
 QY 3650 GAGACATGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3709
 DB GAGACATGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3709
 QY 2194 GAGACATGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 2195
 DB GAGACATGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 2195
 QY 3708 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3765
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3765
 QY 2074 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 2075
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 2075
 QY 3766 ATCAGCCTTTTAAAGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3824
 DB ATCAGCCTTTTAAAGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3824
 QY 2014 CTTCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 2055
 DB CTTCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 2055
 QY 3824 CTTCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3884
 DB CTTCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3884
 QY 1954 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1995
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1995
 QY 3895 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3944
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 3944
 QY 1894 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1936
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1936
 QY 3945 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4004
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4004
 QY 1835 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1876
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1876
 QY 4005 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4064
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4064
 QY 1775 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1776
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1776
 QY 4065 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4114
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4114
 QY 1715 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1756
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1756
 QY 4115 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4174
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4174
 QY 1655 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1696
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1696
 QY 4175 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4228
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4228
 QY 1594 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1636
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1636
 QY 4239 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4285
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4285
 QY 1535 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1576
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1576
 QY 4286 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4337
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4337
 QY 1475 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1516
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1516
 QY 4338 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4397
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4397
 QY 1415 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1456
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1456
 QY 4399 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4455
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4455
 QY 1355 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1396
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1396
 QY 4456 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4515
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 4515
 QY 1295 GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1336
 DB GTCCTGATGATATATAGCCCTTTTTTTTTTTTTTTTATAGATCTCT 1336

XX Nucleic acids encoding human immunodeficiency virus (HIV) antigens and polypeptides, for preventing, diagnosing and/or treating AIDS and related diseases and metastasis.

XX Disclorator: SEQ ID NO 25240; 3071bp - Sequence listing, English.

XX AAK54451 to AAK54702 encode the human immunodeficiency virus (HIV) antigen (i) amino acid sequences given in AAK54703 to AAK54704, (ii) have cytostatic activity, and can be used in gene therapy and vaccine production; (iii) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate HIV expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (i) by expressing inactive proteins or to supplement the patient's own production of (i). Additionally, (ii) polynucleotides may be used to produce the secreted HIV by inserting the nucleic acids into a host cell and culturing the cell to express the protein; (iii) proteins and polynucleotides may be used to prevent, diagnose and treat immunodeficiency related diseases, especially cancers and cancer metastases of hematopoietic derived cells. AAK54703 to AAK54704 represent human immunodeficiency virus (HIV) antigen sequences from the present invention. AAK54705 to AAK54706 and AAK54707 represent sequences used in the exemplification of the present invention. Sequences 26277 BP; 5199 A; 7007 C; 7272 G; 4033 T; 0 other;

DP 51531 STATTCCGACTACTGGGAGCGTGAACAGAGAAATCTTGAACTCCAGAGGACAGG 51532
 QY 4312 TTTTGTGAGTGAAGTGGAGCGATTTACCTTCAGTCTGGCGACACAGAGCGGAATCCG 4371
 PB 51541 TTTGAGTGAAGTGGAGCGATTTACCTTCAGTCTGGCGACACAGAGCGGAATCCG 51472
 QY 4072 TCTCAAAA...AAAAAAGAGGCGGGGTTTCTTCAAGCAAGCGCGCTTTGACAGCAAT 4128
 PB 51471 TCTCAAAATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 51412
 QY 4123 GTTATTCTCGGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4158
 PB 51411 TTTTTCACACAGTCTTTTGGGAGCAAGTAAATAGAGAGTAAATAGAGTAAATAG 51352
 QY 4180 TTTTTCACACAGTCTTTTGGGAGCAAGTAAATAGAGAGTAAATAGAGTAAATAG 4248
 PB 51351 TTTTTCACACAGTCTTTTGGGAGCAAGTAAATAGAGAGTAAATAGAGTAAATAG 51294
 QY 4249 GAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4297
 PB 51291 GAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 51234
 QY 4298 GCGACACTAGCTCTGTAGTACTGTAGTACTGTAGTACTGTAGTACTGTAGTACTGTAG 4357
 PB 51213 GCGACACTAGCTCTGTAGTACTGTAGTACTGTAGTACTGTAGTACTGTAGTACTGTAG 51174
 QY 4358 TTTTACTTTTCTTTTGGAGATTTTGGAGATTTTGGAGATTTTGGAGATTTTGGAGAT 4417
 PB 51171 TTTTACTTTTCTTTTGGAGATTTTGGAGATTTTGGAGATTTTGGAGATTTTGGAGAT 51145
 QY 4418 ATCTCAAAAAAATAAAGAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 4477
 PB 51141 GGTCAATTTTGGCAAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 51085
 QY 4478 ACAACATAGAGCTATCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4537
 PB 51084 TCACTATTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 51025
 QY 4538 TTTTCTCAAAAAAACAACACTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4597
 PB 51024 CTCTCAAAAGTGTGGATTACATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 50965
 QY 4598 TCAAAAAA...TCAACACAGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGG 4657
 PB 50964 CAGGAAATAAAGGAATAAGCGCGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGG 50905
 QY 4658 AGCTGAGCGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGG 4714
 PB 50904 AGCTGAGCGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGG 50845
 QY 4715 GAAACCGCTCTCTACTAAAAATCAAAAAATTTAGTCTGGTCTGGTCTGGTCTGGTCTGGTCTG 4774
 PB 50844 GAAACCTCCATCTCTACTAAAAATCAAAAAATTTAGTCTGGTCTGGTCTGGTCTGGTCTGG 50785
 QY 4775 TCACAGCTACTCAGAGGCTGAGTGAAGATCACTTAACTCCAGAGCGGAGGTTGC 4834
 PB 50784 TCCAGCTACTCAGAGGCTGAGTGAAGATCACTTAACTCCAGAGCGGAGGTTGC 50725
 QY 4835 AGTGAAGGAGATCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4894
 PB 50724 AGTGAAGGAGATCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 50665
 QY 4895 AAAAAA...AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4911
 PB 50664 AAAAAA...AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 50428

RESULT 10

ABL64414/C

ID ABL64414 standard; DNA; 65608 bp.

XX

AC

XX

XX

XX

DT 15-MAY-2002 (first entry)
 DE Stomach cancer related gene sequence SEQ ID NO:2751.
 DE Human; cancer; colon; breast; ovary; esophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumor; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumor; adenocarcinoma;
 KW gene; ds.
 XX Homo sapiens.
 CS WC200194629 AZ.
 PN 13-DEC-2001.
 PO 30-MAY-2001; 2001WC-051049

PR 35-JUN-2002; 2000US-209474P
 PR 35-JUN-2002; 2000US-209531P
 PR 18-SEP-2000; 2000US-211133P
 PR 18-SEP-2000; 2000US-211613P
 PR 20-SEP-2000; 2000US-214209P
 PR 20-SEP-2000; 2000US-214314P
 PR 20-SEP-2000; 2000US-214552P
 PR 22-SEP-2000; 2000US-214509P
 PR 22-SEP-2000; 2000US-214567P
 PR 25-SEP-2000; 2000US-214924P
 PR 25-SEP-2000; 2000US-214924P
 PR 25-SEP-2000; 2000US-214977P
 PR 25-SEP-2000; 2000US-215052P
 PR 25-SEP-2000; 2000US-215134P
 PR 25-SEP-2000; 2000US-215280P
 PR 26-SEP-2000; 2000US-215637P
 PR 26-SEP-2000; 2000US-215637P
 PR 27-SEP-2000; 2000US-215711P
 PR 27-SEP-2000; 2000US-215720P
 PR 27-SEP-2000; 2000US-215840P
 PR 27-SEP-2000; 2000US-215863P
 PR 28-SEP-2000; 2000US-216288P
 PR 28-SEP-2000; 2000US-216332P
 PR 28-SEP-2000; 2000US-216332P
 PR 28-SEP-2000; 2000US-216334P
 PR 28-SEP-2000; 2000US-216399P
 PR 28-SEP-2000; 2000US-216111P
 PR 29-SEP-2000; 2000US-216842P
 PR 29-SEP-2000; 2000US-216891P
 PR 02-OCT-2000; 2000US-217122P
 PR 02-OCT-2000; 2000US-217123P
 PR 02-OCT-2000; 2000US-217288P
 PR 02-OCT-2000; 2000US-217294P
 PR 02-OCT-2000; 2000US-217295P
 PR 02-OCT-2000; 2000US-217316P
 PR 03-OCT-2000; 2000US-217425P
 PR 03-OCT-2000; 2000US-217598P
 PR 03-OCT-2000; 2000US-217604P
 PR 03-OCT-2000; 2000US-217606P
 PR 03-OCT-2000; 2000US-217609P
 PR 01-NOV-2000; 2000US-244867P
 PR 01-NOV-2000; 2000US-245084P

(AVAL-): AVALON PHARM.

Young Ph, Augustus N, Carter KC, Ebner R, Endress G, Horrigan S;
 Sopet DR, Weaver Z;
 WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a
 chemical agent to be tested for anti-neoplastic activity, and
 determining a change in expression of a gene of a signature gene set

Claim 1; SEQ ID 2751; 44pp; English.

QY 1494 TGTGGAGGAGAACGGCCACAGGAGCTGTTCTTACAGGCTTATTGGAGAGAGCTG 1553
 DE 1093 GACTGGAGAGGGGGAGAGAGATTTCTGTAGCGAGCTTCATCAGAAAGAGCTG 1152
 QY 1554 CTCCCTGAGCTATCTTGTCTCATCACCACAGGAGGAGCTTTTGGAGAGCTCCACCGT 1613
 DB 1153 CTCCGAGGCTTCTGCTCATCACCACAGGAGGAGCTTTTGGAGAGCTCCACCGT 1613
 QY 1614 CTGCTGGAGAGCCAGGAGCTGAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAA 1673
 DB 1213 TTGTGTGAGAGCTGAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAA 1673
 QY 1674 TACTTCTACAGTATTTCCCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1733
 DB 1273 TACTTCTACAGTATTTCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1733
 QY 1734 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 1793
 DB 1313 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 1793
 QY 1794 TGCCTCCAGAGAGCTGAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 1853
 DB 1333 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 1853
 QY 1854 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 1913
 DB 1413 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 1913
 QY 1914 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 1973
 DB 1513 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 1973
 QY 1974 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 2033
 DB 1613 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 2033
 QY 2034 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 2093
 DB 1633 GAGAGAGAGCTTCTTACAGAGCTGAGAGCTTCTGTAGAGAGAGAGAGAGAGAGAG 2093

RESULT 15
 ID AEX93556 standard; CCNA; 3857 BP.
 AC AEX93556;
 CT 22-MAY-2003 (first entry)
 CE Huma cDNA encoding PYRN-1.
 KW Human; ss: Gene: nucleotide binding site; pyrin domain; NBS-1; PYRN-1;
 KW caspase 1; anti-inflammatory; apoptosis; ASC; NF-kB; nuclear factor KB;
 KW LRR; leucine rich repeat; inflammatory disorder; familial cold urticaria;
 KW arthritis; inflammatory bowel disease; Crohn's disease; asthma; allergy;
 KW ulcerative colitis; rheumatoid arthritis; Lyme disease; Graves disease;
 KW insulin-dependent diabetes; multiple sclerosis; contact dermatitis;
 KW psoriasis; graft versus host disease; food allergy; conjunctivitis;
 KW chronic obstructive pulmonary disease; food allergy;
 KW helminthic infection; leishmaniasis; viral infection; HIV infection;
 KW bacterial infection; tuberculosis; leprosy; tuberculosis 1q44.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT 139..3243
 FT /tag= a
 FT /product= "PYRN-1"
 XX GS2002187922-A1.
 XX 12-DEC-2002.

XX 22-APR-2002; 2002US 0137516.
 PR 17-FEB-2000; 2000US 0556067.
 PR 01-SEP-2000; 2000US 0653901.
 PR 26-SEP-2001; 2001US 094995.
 PR 20-DEC-2001; 2001US 0227629.
 PA (BERT/) BERTIN J.
 PA (MANU/) MANU: G A.
 XX Bertin J, Manu G A.
 XX WP1: 2003 328763/11.
 DR P-PSDB; ABL048503.
 XX Identification of compounds that binds to polypeptide (for e.g. a PYRN protein), useful for treating disorders associated with inappropriate apoptosis, for e.g. inflammatory disorder.
 XX Disclosure; Fig 4: 74pp; English.
 XX The invention relates to a new method for the identification of a compound that binds to a human PYRN-1 (an apoptotic signaling molecule appearing as ABL0503) comprising contacting the polypeptide or a cell expressing the polypeptide to a test compound and determining whether PYRN-1 binds to the test compound, or determining the effect of the test compound on the activity of PYRN-1, in a similar manner.
 XX Compounds are isolated which modulate the binding of PYRN-1 to ASC comprising testing the compounds against the pyrin binding domains of both PYRN-1 and ASC (not defined), modulate the activity of NF-kB (nuclear factor KB), and which modulate the ASC-mediated activation of NF-kB (comprising: (a) measuring the binding of a test compound to the LRR (leucine rich repeat) domain of PYRN-1; and (b) measuring the activation of NF-kB in a cell expressing ASC and PYRN-1 in the presence and absence of the LRR domain binding compound). The candidate modulators may be useful for treating an inflammatory disorder.
 XX Also included is a method (ME) for identifying a modulator of caspase-1 activity. The identified compound can be used to modulate the polypeptide's activity (ASC and NF-kB activities in a patient). By modulating the expression or activity of the polypeptide (PYRN-1) and ASC, a disorder associated with inappropriate apoptosis (inflammatory disorders e.g. familial cold urticaria, arthritis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, rheumatoid arthritis, Lyme disease, insulin-dependent diabetes, multiple sclerosis, Graves disease, contact dermatitis, psoriasis, graft versus host disease, asthma, chronic obstructive pulmonary disease, allergies (e.g. food allergies), conjunctivitis, helminthic infection (e.g. leishmaniasis), viral infections such as HIV infection, and bacterial infections such as tuberculosis and leprosy) can be treated. Human NBS-1 protein (nucleotide binding site protein 1) contains a pyrin domain which was used to identify genomic sequences containing the PYRN-1 gene. The gene for PYRN-1 is located on chromosome 1q44. The present sequence encodes human PYRN-1.
 XX Sequence 3857 BP; 958 A; 931 C; 1023 G; 945 T; 0 other;
 QY Query Match 7.6%; Score 375.2; DB 25; Length 3857;
 DB Best Local Similarity 60.2%; Pred. No. 7.7e-64;
 QY Matches 719; Conservative 0; Mismatches 379; Indels 97; Gaps 2;
 DB 985 TGGCTCTGGAGTCTTCTGTCTCTCCAGAGAGAGATCCAGGAGAACCTACAGGAC 1044
 DB 499 TACCTTCGAGATCTCTCTTGTAAATGAAGAGAGATTACCGTAGAGATACAGAA 558
 QY 1045 TATGTCCGAGGAGATTCGGGCTCATGAGACCGCATCGCCCTAGGGGAATGTGTC 1104
 DB 559 TACGTGAGAGAGAGATTCAGTCCATTCAGACAGAGATCCCGTGTGAGAGTGT 618
 QY 1105 AATGTCAAGGACCGGTACACCGGCTCTGTGTGTGAGAGAGACTCAACCCCATCGAG 1164
 DB 619 AGCCTCAAGAAACCGGTACACAGACTGCTCTCATCAGGAGAGACCGGAGTCAGAG 678

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

GM nucleic - nucleic search, using sw model

Run on: October 30, 2003, 21:38:28, Session time 17529 Seconds
(without alignment) 11508.093 Million cell updates/sec

Title: US-10-028-374-1

Perfect score: 4931

Sequence: 1 ccacgcgcgcgaacatgggc.....dadaaaaaaaaaa 4931

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2893711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 577422

Minimum DB seq length: 9

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

GenEmbl:

1: gb_ba:

2: gb_hg:

3: gb_in:

4: gb_or:

5: gb_ov:

6: gb_pat:

7: gb_ph:

8: gb_pi:

9: gb_pr:

10: gb_ro:

11: gb_sta:

12: gb_sy:

13: gb_un:

14: gb_vl:

15: gb_ba:

16: gb_in:

17: gb_or:

18: gb_ov:

19: gb_pat:

20: gb_ph:

21: gb_pi:

22: gb_pr:

23: gb_ro:

24: gb_sta:

25: gb_sy:

26: gb_un:

27: gb_vl:

28: gb_ba:

29: gb_in:

30: gb_or:

31: gb_ov:

32: gb_pat:

33: gb_ph:

34: gb_pi:

35: gb_pr:

36: gb_ro:

37: gb_sta:

38: gb_sy:

39: gb_un:

40: gb_vl:

41: gb_ba:

42: gb_in:

43: gb_or:

44: gb_ov:

45: gb_pat:

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	4931	100.0	4931	6	AX684291	Sequence
2	1733.2	34.7	3827	9	AY095146	Homo sapi
3	1733.2	34.7	3827	9	AY154467	Homo sapi
4	1733.2	34.7	3827	9	AY154467	Homo sapi
5	1710	34.7	3507	9	BC028069	Homo sapi
6	1697	34.4	3543	9	AY116205	Homo sapi
7	1697	34.4	3543	9	AY116205	Homo sapi
8	1616	32.8	3128	6	AX417214	Sequence
9	1616	32.8	3128	6	AX417214	Sequence
10	1577.6	32.0	3100	6	AX459869	Sequence
11	1524.8	30.9	3121	9	AY116206	Homo sapi
12	1226.2	24.9	14730	9	AC008753	Homo sapi
13	826	16.8	2158	6	AX575503	Sequence
14	783.6	15.9	2494	9	AF231021	Homo sapi
15	723.8	14.5	22344	2	AC104193	Rattus no
16	662.6	13.4	21827	2	AC079499	Mus muscu
17	507.4	10.3	98758	9	HS223H9	Human DNA
18	460.8	9.3	19870	9	AC112215	Human DNA
19	457.6	9.3	186734	9	AC104446	Homo sapi
20	457.6	9.3	215929	9	AC105934	Homo sapi
21	454.6	9.2	13592	9	AC167162	Human DNA
22	452.8	9.2	59936	9	AC29592	Human DNA
23	452.8	9.2	180283	9	AF134726	Homo sapi
24	451.8	9.2	20724	9	BX284678	Human DNA
25	446.2	9.0	42265	9	AC010504	Homo sapi
26	446.2	9.0	179934	9	AL662834	Human DNA
27	445.4	9.0	92766	2	AF322454	Homo sapi
28	445.4	9.0	101685	9	HS019622	Human DNA
29	435	8.8	195952	9	AC212184	Homo sapi
30	430.2	8.7	121332	9	AC006441	Homo sapi
31	429.6	8.7	165617	9	AF165926	Homo sapi
32	429.6	8.7	177782	9	AC117532	Homo sapi
33	429.4	8.7	116762	9	AC108004	Homo sapi
34	428	8.7	107717	9	AC132555	Homo sapi
35	428	8.7	108139	2	AC004285	Continuation of
36	427.6	8.7	100000	9	AF000535	Homo sapi
37	426	8.6	10127	9	AC104134	Homo sapi
38	426	8.6	160335	9	AB000729	Homo sapi
39	426	8.6	177592	9	AC093814	Homo sapi
40	426	8.6	162769	9	AC003176	Homo sapi
41	426	8.6	189406	9	AC112717	Homo sapi
42	425.2	8.6	13892	2	HS112K5	Human DNA s
43	422.2	8.6	131386	2	AC137962	Homo sapi
44	421	8.5	137786	9	AL450126	Human DNA
45	420.4	8.5	196686	9	AC010128	Homo sapi

ALIGNMENTS

RESULT:

AX684291

LOCUS

DEFINITION

Sequence 1 from Patent WO2005011

AX684291

ACCESSION

VERSION

AX684291.1

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Eutelesstomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

Feder, J., Ramakrishnan, C. and Mintier, S.

Human leucine rich repeat containing protein, Lrrbml, expressed

predominately in bone marrow

PAT 29-MAR-2003

Pred. No. is the number of results predicted by chance to have a

661 TCTTTGTTGTAAGTGTTGTTCTGCTGGAAATATTTCTCCATCTCTTTTGCTGGAAAGATA 720

721 TACTTACTGAAGCTTTGAAGCACTACGTTCTCTCTTAAACATTGCAAAATGCAATTAATTTG 780

722 TACTTACTGAAAGCTTGAAGCACTACGTTCTCTCTTAAACATTGCAAAATGCAATTAATTTG 780

781 CCGAGCGCTTTTCGCGGAACTTCATGCTTTAGTCTTACTGAGAGCACTGACAGTTCTCC 840

781 CCGAGCGCTTTTCGCGGAACTTCATGCTTTAGTCTTACTGAGAGCACTGACAGTTCTCC 840

841 AATGAGACATATTTGAGATTTGCTCTGCTGCTCTGTTGATGCTGAGAAATGATATCC 900

841 AATGAGACATATTTGAGATTTGCTCTGCTGCTCTGTTGATGCTGAGAAATGATATCC 900

901 CCGTGGAGATTTATCTCTATCTGTTAAACAGAGGATCAGATATCCCACTCTGTTGCGCGG 960

901 CCGTGGAGATTTATCTCTATCTGTTAAACAGAGGATCAGATATCCCACTCTGTTGCGCGG 960

961 TCTTCACTTTGAAATGAGTGAATATGCTTTCTGGAAATCTCTCTGTCTACCTCCAAAGAAA 1020

961 TCTTCACTTTGAAATGAGTGAATATGCTTTCTGGAAATCTCTCTGTCTACCTCCAAAGAAA 1020

1021 GATCCCGAGAAATCTTAAAGGACATATCTCCGACGAAATTCGGGCTCATGGAACACCGC 1080

1021 GATCCCGAGAAATCTTAAAGGACATATCTCTCCGAGAAATTCGGGCTCATGGAACACCGC 1080

1081 AATGCGGCTATGAGGAAATGTTCTGAACCTTCAAGCAAGGATCAAGCGGCTCTGCTGCTG 1140

1081 AATGCGGCTATGAGGAAATGTTCTGAACCTTCAAGCAAGGATCAAGCGGCTCTGCTGCTG 1140

1141 AAGTAGCAATCAAAAGTCAATGAGTCAAGCAGCAGCTCTGACACAGCGCGGGACAC 1200

1141 AAGTAGCAATCAAAAGTCAATGAGTCAAGCAGCAGCTCTGACACAGCGCGGGACAC 1200

1201 GCGAGGAGATTCGGAACACAGGTAGCTGATCAAGATAGAGACCTCTTTGAGCGACAC 1260

1201 GCGAGGAGATTCGGAACACAGGTAGCTGATCAAGATAGAGACCTCTTTGAGCGACAC 1260

1261 GAGAGAGCGCTCCAGATATCCGCTCATCGTGTATCGTAAGCGCGGACGGGATAGGAGA 1320

1261 GAGAGAGCGCTCCAGATATCCGCTCATCGTGTATCGTAAGCGCGGACGGGATAGGAGA 1320

1321 TGAACACAGATGTCTGGAATTCAGATATCAAGACCTCATCTTACGCTCTGATCTGAGC 1380

1321 TGAACACAGATGTCTGGAATTCAGATATCAAGACCTCATCTTACGCTCTGATCTGAGC 1380

1381 CCAAGAGCGCTTCAAGAGCCTATCGAGTTCGCGAGCGGCTCTTTCATCATCGACG 1440

1381 CCAAGAGCGCTTCAAGAGCCTATCGAGTTCGCGAGCGGCTCTTTCATCATCGACG 1440

1441 GATTCGATATATTAAGATTTTTCGAGATTCGTAAGGACCTTGTTGCTCTGCTGCTG 1500

1441 GATTCGATATATTAAGATTTTTCGAGATTCGTAAGGACCTTGTTGCTCTGCTGCTG 1500

1501 AGGAGAAAGCGCCATCAAGAGCTGCTCTTTAAACAGCTTAATTCGGAAGAGCTGCTCTCTG 1560

1501 AGGAGAAAGCGCCATCAAGAGCTGCTCTTTAAACAGCTTAATTCGGAAGAGCTGCTCTCTG 1560

1561 AGCTATCTTCTCATCAACACAGCGCCACCGCTTTCGAGAGCTCCACCGCTCTGCTGG 1620

1561 AGCTATCTTCTCATCAACACAGCGCCACCGCTTTCGAGAGCTCCACCGCTCTGCTGG 1620

1621 AATACCTTCAAGATGTAAGATTCCTAGGCTCTTCTGAGGGCAGAAAGAGGAATACTTCT 1680

1621 AATACCTTCAAGATGTAAGATTCCTAGGCTCTTCTGAGGGCAGAAAGAGGAATACTTCT 1680

1681 ACAAGTATTTCAAGATTCAGAGACAGCGGCGCAAGCTTTTCAATTACCTGAGGACACACG 1740

1681 ACAAGTATTTCAAGATTCAGAGACAGCGGCGCAAGCTTTTCAATTACCTGAGGACACACG 1740

1741 AATTTCTTTTACGATGCTTCTGCTGCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800

1741 AATTTCTTTTACGATGCTTCTGCTGCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800

1801 AGCAGCAGCTGGAGGGTGGGGGCTCTTCAGACAGACGTTCAGGACCACTGCAGTGT 1860
1801 AGCAGCAGCTGGAGGGTGGGGGCTCTTCAGACAGACGTTCAGGACCACTGCAGTGT 1860
1961 ACATGCTCTACCTGCTGAGTCTGATCAACCGAAGCTGGGGGCTCGGCTCTCCAGCCCC 1920
1961 ACATGCTCTACCTGCTGAGTCTGATCAACCGAAGCTGGGGGCTCGGCTCTCCAGCCCC 1920
1921 CACCCAAACAGAGAGGTTGTGCTCTCTGGCGGAGATGGGCTCTGGAAATCAGAAATCC 1980
1921 CACCCAAACAGAGAGGTTGTGCTCTCTGGCGGAGATGGGCTCTGGAAATCAGAAATCC 1980
1981 TATTTCAGGAGCAGGACCTCCGGAAGCAGCGCTAGACGCGGAAGACGTCTCTGCTTCC 2040
1981 TATTTCAGGAGCAGGACCTCCGGAAGCAGCGCTAGACGCGGAAGACGTCTCTGCTTCC 2040
2041 TCACATGAACATCTTCAGAGAGGATTAACATGTGTAGAGAGGCTTCTGGCAATCAGCA 2100
2041 TCACATGAACATCTTCAGAGAGGATTAACATGTGTAGAGAGGCTTCTGGCAATCAGCA 2100
2101 GCGGCTTCTCTGTTGGACTCTTAAGAGAGAGACAGGAGCTACCTGAGAGAGCTCTCT 2160
2101 GCGGCTTCTCTGTTGGACTCTTAAGAGAGAGACAGGAGCTACCTGAGAGAGCTCTCT 2160
2161 GCTGGAGGTTCTCGCGGACATCAAGATGGAAGCTCTGAGTGGATCCAAAGCAAGCTC 2220
2161 GCTGGAGGTTCTCGCGGACATCAAGATGGAAGCTCTGAGTGGATCCAAAGCAAGCTC 2220
2221 AGACCGACGCTCGACCTCGACAGAGGCTCTCTGAGTGGATCCAAAGCAAGCTC 2280
2221 AGACCGACGCTCGACCTCGACAGAGGCTCTCTGAGTGGATCCAAAGCAAGCTC 2280
2281 TCACAGAGGAGGAGTATATCCAGACAGGCTCTGAGTGGATCCAAAGCAAGCTC 2340
2281 TCACAGAGGAGGAGTATATCCAGACAGGCTCTGAGTGGATCCAAAGCAAGCTC 2340
2341 ACATTGCTCTCAAGATGGAGCAGATGCTCTCTGAGTGGATCCAAAGCAAGCTC 2400
2341 ACATTGCTCTCAAGATGGAGCAGATGCTCTCTGAGTGGATCCAAAGCAAGCTC 2400
2401 CCGAGTGTGCTGCTGTAAGGCGCAGCTACAGTGGAGGAGGAGGAGGAGGAGGAGG 2460
2401 CCGAGTGTGCTGCTGTAAGGCGCAGCTACAGTGGAGGAGGAGGAGGAGGAGGAGG 2460
2461 GCTCGCAGAGGCGCACACCTGCTGAGTACAGAGAGGAGGAGGAGGAGGAGGAGG 2520
2461 GCTCGCAGAGGCGCACACCTGCTGAGTACAGAGAGGAGGAGGAGGAGGAGGAGG 2520
2521 CTTACAGTGAACATCTGGCAGGCGCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 2580
2521 CTTACAGTGAACATCTGGCAGGCGCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 2580
2581 TGTACCGAATGCTGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2640
2581 TGTACCGAATGCTGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2640
2641 CCAACTGCAAACTCAGAACTGAGTAAATTTATATATATAAATATATATATATATAT 2700
2641 CCAACTGCAAACTCAGAACTGAGTAAATTTATATATATAAATATATATATATATAT 2700
2701 TAAATATATGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
2701 TAAATATATGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
2761 GATGGGAGGATCATCTTACCCAGGAGTTTAAAGCAGGAGTGGCTTAACTATGATGAA 2820
2761 GATGGGAGGATCATCTTACCCAGGAGTTTAAAGCAGGAGTGGCTTAACTATGATGAA 2820
2821 CATCTCTACTTAAATACCAAAATGAGCGAGGATGAGCGAGGATGAGCGAGGATGAG 2880
2821 CATCTCTACTTAAATACCAAAATGAGCGAGGATGAGCGAGGATGAGCGAGGATGAG 2880
2821 CATCTCTACTTAAATACCAAAATGAGCGAGGATGAGCGAGGATGAGCGAGGATGAG 2880

2881 TACTCAGGAGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940
2881 TACTCAGGAGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940
2941 GAGTGCCGCTCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 3000
2941 GAGTGCCGCTCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 3000
3001 GATTTTGAAGAGATGAGTCTCAGTGGCAAGGCGCTGGATTCAGGAGGAGGAGGAG 3060
3001 GATTTTGAAGAGATGAGTCTCAGTGGCAAGGCGCTGGATTCAGGAGGAGGAGGAG 3060
3061 TTTGAGGAGGCTCTGCGGCTATCCCAATGCGAGGCTGAGATGATTCAGTTSAGGAG 3120
3061 TTTGAGGAGGCTCTGCGGCTATCCCAATGCGAGGCTGAGATGATTCAGTTSAGGAG 3120
3121 GCTTGAATCTCGGAGGCTCTCAGGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3180
3121 GCTTGAATCTCGGAGGCTCTCAGGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3180
3181 TGAATTTGAGAGATGAGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240
3181 TGAATTTGAGAGATGAGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240
3241 ACTGAGGAGGCTCTGCGGCTATCCCAATGCGAGGCTGAGATGATTCAGTTSAGGAG 3300
3241 ACTGAGGAGGCTCTGCGGCTATCCCAATGCGAGGCTGAGATGATTCAGTTSAGGAG 3300
3301 TGTGCTCTGAGAGGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3360
3301 TGTGCTCTGAGAGGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3360
3361 CTTGAGGAGGCTCTGCGGCTATCCCAATGCGAGGCTGAGATGATTCAGTTSAGGAG 3420
3361 CTTGAGGAGGCTCTGCGGCTATCCCAATGCGAGGCTGAGATGATTCAGTTSAGGAG 3420
3421 TGTGCTCTGAGAGGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480
3421 TGTGCTCTGAGAGGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480
3481 CCAATTCATGAGAGGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3540
3481 CCAATTCATGAGAGGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3540
3541 AATCTTACTATCTAGCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3600
3541 AATCTTACTATCTAGCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3600
3601 ATTACAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
3601 ATTACAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
3661 TCTATCATGTTAGCAGGATGCTCTGATCTACTACTACTACTACTACTACTACTACTACT 3720
3661 TCTATCATGTTAGCAGGATGCTCTGATCTACTACTACTACTACTACTACTACTACTACT 3720
3721 CCAAGTCTCTGAGATTTAGAGGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGG 3780
3721 CCAAGTCTCTGAGATTTAGAGGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGG 3780
3781 AGGATTTTCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3840
3781 AGGATTTTCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3840
3841 AGTGCGAGGATCTACTGAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3900
3841 AGTGCGAGGATCTACTGAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3900
3901 CCGATCT 3960
3901 CCGATCT 3960
3961 GCTACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4020


```
DB      645  TACATCCGGCTCTGCTGGTGAAGAGAGACACTCAAACTCCAATGAGGTCCACAGACAGCTT 704
CY      1:80  CTGACACAGAGCCGGGACACAGGAGACAGTGGAGACAGCAAGCTAGCCGCTCAAGATA 1239
DB      925  CTGGACACAGCCGGGACACAGGAGACAGTGGAGACAGCAAGCTAGCCGCTCAAGATA 764
CY      1240  GAGACCTCTTTGAGCCAGACAGAGAGAGCCGCGAGACAGAGGACAGGCTGATCAAA 1299
DB      765  GAGACCTCTTTGAGCCAGACAGAGAGAGAGGCTGATCAAGAGAGAGGCTGATCAAA 824
CY      1300  GAGCGGCGAGGAGT... 1313
DB      825  GCGCGCGGAGGATAGGCAATCCAGATGAGATGAGACAGAGAGAGGCTGATCAAGAGAG 884
CY      1314  ..... 1325
DB      885  GAGAGAGCTTTCCAGAGGAGATTTTATATATTTTATATCAACTCAGGAGATGAAC 944
CY      1326  CAGAGTGCACGAAATGAGCATGACAGACCTCACTTCAATATATGCTGCTGAGCCAGC 1385
DB      945  CAGAGTGCACGAAATGAGCATGACAGCTCACTTCAATATATGCTGCTGAGCCAGC 1004
CY      1386  GCGCTCTCCAGAGACTCATGAGATTCGCGAGAGAGGCTTTCATCATATGAGAGGCTTC 1445
DB      1005  GCGCTCTCCAGAGACTCATGAGATTCGCGAGAGAGGCTTTCATCATATGAGAGGCTTC 1564
CY      1446  GATGAGTCAAGACCTTTCTTCAGCATTCGAGAGAGAGGCTTTCATCATATGAGAGG 1505
DB      1045  GATGAGTCAAGACCTTTCTTCAGCATTCGAGAGAGAGGCTTTCATCATATGAGAGG 1124
CY      1546  AAACGGCCACAGAGACTCTTTTAAACAGCTTAAATTTGAGAGAGAGGCTTTCATTA 1565
DB      1125  AAACGGCCACAGAGACTCTTTTAAACAGCTTAAATTTGAGAGAGAGGCTTTCATTA 1184
CY      1566  TCTTTGCTGATCACCAGAGGCTTTCAGAGGCTTTCAGAGAGAGGCTTTCAGAGGCT 1625
DB      1185  TCTTTGCTGATCACCAGAGGCTTTCAGAGGCTTTCAGAGAGAGGCTTTCAGAGGCT 1244
CY      1626  CCGAGGATGCGAGACTCTGAGCTTTCAGAGGCTTTCAGAGAGAGGCTTTCAGAGG 1685
DB      1245  CCGAGGATGCGAGACTCTGAGCTTTCAGAGGCTTTCAGAGAGAGGCTTTCAGAGG 1304
CY      1686  TATTTCCAAATGCGAGAGAGGCTTTCAGAGGCTTTCAGAGAGAGGCTTTCAGAGG 1745
DB      1305  TATTTCCAAATGCGAGAGAGGCTTTCAGAGGCTTTCAGAGAGAGGCTTTCAGAGG 1364
CY      1746  CTCTTCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1805
DB      1365  CTCTTCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1424
CY      1806  CAGCTGGAGGCTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1865
DB      1425  CAGCTGGAGGCTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1484
CY      1866  CTCTACCTGCTGAGTCTGATGCAACCTGCAAGAGAGGCTTTCAGAGGCTTTCAGAG 1925
DB      1485  CTCTACCTGCTGAGTCTGATGCAACCTGCAAGAGAGGCTTTCAGAGGCTTTCAGAG 1544
CY      1926  AACCCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1985
DB      1545  AACCCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1604
CY      1986  GAGGAGCAGAGCTCCGGAGACAGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCT 2045
DB      1605  GAGGAGCAGAGCTCCGGAGACAGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCT 1664
CY      2046  ATGACATCTTCAGAGAGAGATCACTGCTGAG 2078
DB      1665  ATGACATCTTCAGAGAGAGATCACTGCTGAGAGAGATCACTGCTGAGAGAGATCA 1724
CY      2079  ..... 2078
```


[illegible]

Db	2385	CTCAGACACCCCAACTGCGAACTTTCAGACCTCA	2418
QY	2691	TATTTTAAATAATATATTGCCCAAGTATGATGGCTCAGCGCTGTAAATTCACGCACTTT	2750
Db	2419	2418
QY	2751	GGGAGCCACAGATGGGAGGATCACTTTCACACAGGAGTTCAAGACCAAGCTGCGCAACAAT	2810
Db	2419	2418
QY	2811	GCTGAATTCCTATCTCTTACTAATAAATATTAATAATGACCCAGGCATGCTGCACACAGCTCTG	2870
Db	2419	2418
QY	2871	TACACCTCACTCTTCAGACAGGCAAGAGGAGGATTCCTTCAACCCAGAGGCGCAGGCT	2930
Db	2419	2418
QY	2931	TGTGGCTTAAGAGTGTCTGCATCTCCAGCTCAGCTCCGAGGAGCTCTCTCAGCTCTCA	2990
Db	2419	2418
QY	2991	TAGCCCATTAAGATTTTGAACAGATCGATCTCAGTGGCAACGGCGTTGGATTCCTCAGCA	3050
Db	2476	TACCCATTAAGAAATTTGAAAGATGATCTCTCAGTGGCAACGGCGTTGGATTCCTCAGCA	2515
QY	3051	TGATCGCTCTTCGAGGAGGCTCGGCACTCCCAATGCAAGCTCGAGCTGATTTAGTTGA	3110
Db	2536	TGATCGCTCTTCGAGGAGGCTCGGCACTCCCAATGCAAGCTCGAGCTGATTTAGTTGA	2595
QY	3111	GGAGGTGTACGTGGAGTTCGGGCTCTGCAGAGATGGCTCTGTGCTTGGCAACCAAC	3170
Db	2596	GGAGGTGTACGTGGAGTTCGGGCTCTGCAGAGATGGCTCTGTGCTTGGCAACCAAC	2655
QY	3171	CACACTTCGTGAGTTGGACCTCACAGGAAATGCACCTGGAGGATTTGGGCTCGAGTTAC	3230
Db	2656	CACACTTCGTGAGTTGGACCTCACAGGAAATGCACCTGGAGGATTTGGGCTCGAGTTAC	2715
QY	3231	TATGCCAGGACCTTAGGCACTTCTGCAGACTAGGACCTTTGTGGCTGAAGATCTGCC	3290
Db	2716	TATGCCAGGACCTTAGGCACTTCTGCAGACTAGGACCTTTGTGGCTGAAGATCTGCC	2775
QY	3291	GGCTCATCTGTCTGTCTGTACAGCTGGCTGCACTCTCACTGTGACAGAGAGCTGA	3350
Db	2776	GGCTCATCTGTCTGTCTGTACAGCTGGCTGCACTCTCACTGTGACAGAGAGCTGA	2815
QY	3351	GAGAGCTGTACCTGAGCTGAAATGAGCTGGGAGCTCGGGGTGCTGCTCTGTGTGAGG	3410
Db	2836	GAGAGCTGTACCTGAGCTGAAATGAGCTGGGAGCTCGGGGTGCTGCTCTGTGTGAGG	2895
QY	3411	GCTCATCATCTGTCTGTCTGTACAGCTGGCTGCACTCTCACTGTGACAGAGAGCTGA	3474
DE	2896	GCTCATCATCTGTCTGTCTGTACAGCTGGCTGCACTCTCACTGTGACAGAGAGCTGA	2939
RESULT 4			
AK095460			
LOCUS	AK095460	3466 bp	rRNA linear PRI 15-JUL 2002
DEFINITION	Homo sapiens cDNA FLJ18141 fis, clone D90ST2302673, weakly similar to Homo sap:ens caspase recruitment domain protein 7 mRNA.		
ACCESSION	AK095460		
VERSION	AK095460.1	GI:21754727	
KEYWORDS	c.igo capping; fis (full insert sequence).		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokota, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Onuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamaoto, T., Isono, Y., Kawai, H., Y., Saito, K., Nishikawa, T.,		

9.3 CASAGTGCACGGAAATGCAGCA.TGCAGAGC.CATCT.CAGCT.GCTGGCCTAGGCCAGC 972

[illegible]

Db	2864	GCCTCAGGATCCACCGTTCAGACCTTCAGACCTCGCGGTGGAGTGTGTGGCTCACAG	292
Qy	3471	GCATAGS 3476	
Db	2924	CCAGS 2929	
RESULT 5			
LOCUS	BC028069		
DEFINITION	Homo sapiens neuronal apoptosis inhibitor protein 12, mRNA (cDNA clone M3:4117 IMAGE52127); complete cds		
ACCESSION	BC028069.1		
VERSION	1		
KEYWORDS			
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			

[illegible][illegible]

[illegible][illegible]

1659	TTTGGACTCTTGAACGAGGAGACACAGGAGCCACCTGGAGAAGAGTCTCTCTCTGGAAAGTCT	1668
2172	TGCGGCACATCAAGATGGACCTGTTTGCAGTGGATCCAAAGCAAGCTCAGAGCAGCGG	2231
1669	TGCGGCACATCAAGATGGACCTGTTTGCAGTGGATCCAAAGCAAGCTCAGAGCAGCGG	1725
2212	TCCACCTGACAGAGGAGGCTCTTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2291
1729	TCCACCTGACAGAGGAGGCTCTTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	1788
2232	CAGTTTATCCAGAGGAGGCTCTTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2351
1789	GAGTTTATCCAGAGGAGGCTCTTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	1848
2352	AAGATGAGGAGGAGGCTCTTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2411
1849	AAGATGAGGAGGAGGCTCTTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	1908
2412	CAGTTTATCCAGAGGAGGCTCTTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2471
1949	CAGTTTATCCAGAGGAGGCTCTTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	1968
2472	GCGCACACGCTGTTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2528
1969	GCGCACACGCTGTTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2028
2529	GAACATCTGGCAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2588
2029	GAACATCTGGCAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2088
2589	AATGGCTGGCAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2648
2089	AATGGCTGGCAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2148
2649	AAATTTGAGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2708
2149	AAATTTGAGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2164
2709	TGGGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2768
2165	TGGGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2164
2769	GGATGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2828
2165	GGATGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2164
2829	CTAAGATGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2888
2165	CTAAGATGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2164
2889	AGATGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2948
2165	AGATGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2179
2949	GCATCTGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	3008
2180	GCATCTGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2239
3009	CAAGATGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	3068
2240	CAAGATGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2299
3069	GCATCTGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	3128
2300	GCATCTGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2359
3129	GCATCTGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	3186
2360	GCATCTGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2419
3189	GCATCTGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	3248
2420	GCATCTGAGGAGGAGGCTCTTGGTGGAGTCTTTCAGTGTCTTGTACGAGATCCAGAGGAG	2479

Db 223 ACCAGGACCCACCTGGAGAGAGCTCTGTGTGAAGGTCCTGGCGZACATCAGATGGAC 282
 QY CTGTGTGAGTGGATCCAAAGCAAGCTCAGAGTACAGGCTCCACCGCTCAGCAGGCTCC 2252
 Db 284 CTGTGTGAGTGGATCCAAAGCAAGCTCAGAGTACAGGCTCCACCGCTCAGCAGGCTCC 342
 QY TTGAGTCTCTTCAGTCTCTGTATAGATCCAGAGAGAGGATTTATCCAGCAGGCTCC 2312
 Db 342 TTGAGTCTCTTCAGTCTCTGTATAGATCCAGAGAGAGGATTTATCCAGCAGGCTCC 402
 QY AGCCACCTTCAGGTTAGT 2372
 Db 403 AGCCACCTTCAGGTTAGT 462
 QY TGT 2432
 Db 463 TGT 522
 QY AGCCGGGAGGGGAGAGCCGGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2492
 Db 523 AGCCGGGAGGGGAGAGCCGGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 582
 QY CTACAGAGAGAGAGGCTGT 2552
 Db 583 CTACAGAGAGAGAGGCTGT 642
 QY ACCAATCCAAACCTGTATAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2612
 Db 643 ACCAATCCAAACCTGTATAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 702
 QY AAGTGTCTCTTCAGGCTCAG 2672
 Db 703 AAGTGTCTCTTCAGGCTCAG 754
 QY TTATCATATATAACATGATATTATTAATTAATATATGTGTGTGTGTGTGTGTGTGTGTGTGT 2732
 Db 755 TTATCATATATAACATGATATTATTAATTAATATATGTGTGTGTGTGTGTGTGTGTGTGTGT 754
 QY CTGTAAATCCAGAGCTTTGAG 2792
 Db 755 CTGTAAATCCAGAGCTTTGAG 754
 QY GACGAGCTGCGCAACATAGCTCAAACTGATCTCTTAATAAATACCAAAATGAGCAGG 2852
 Db 755 GACGAGCTGCGCAACATAGCTCAAACTGATCTCTTAATAAATACCAAAATGAGCAGG 754
 QY CATGTGTGCACACCTGTCTAAGCTTATATATATATATATATATATATATATATATATATAT 2912
 Db 755 CATGTGTGCACACCTGTCTAAGCTTATATATATATATATATATATATATATATATATATAT 754
 QY AACCCAG 2972
 Db 755 AACCCAG 2972
 QY ACCCTCTGAGCTCTCTATACCAATAGAGATTTATATATATATATATATATATATATATATAT 3032
 Db 754 ACCCTCTGAGCTCTCTATACCAATAGAGATTTATATATATATATATATATATATATATATAT 853
 QY GCGTTGATATCCAGGCTAT 3092
 Db 854 GCGTTGATATCCAGGCTAT 913
 QY TGCACATGATTCAGTTAGGAG 3152
 Db 914 TGCACATGATTCAGTTAGGAG 973
 QY CTGTCTTGGCAGCAACCCACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3212
 Db 974 CTGTCTTGGCAGCAACCCACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1033
 QY ATTTGGGCTGTAGGTTACTATGCAAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3272

Db 1034 ATTTGGGCTGTAGGTTACTATGCAAGAGAGTGAAGGACCCAGCTCTGAGACTACGGACTT 1093
 QY TGTGCTGAGAGATTTGCGGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3332
 Db 1094 TGTGCTGAGAGATTTGCGGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153
 QY CTGTGAGAGCAAGCTCTGAG 3392
 Db 1154 CTGTGAGAGCAAGCTCTGAG 1213
 QY TGTGCTGAGAGATTTGCGGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3452
 Db 1214 TGTGCTGAGAGATTTGCGGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
 QY 3453 TC 3454
 Db 1274 GC 1275

RESULT 14
 AF231021
 LOCUS Homo sapiens leucine rich-repeat protein RNO2 mRNA, complete cds.
 DEFINITION AF231021
 VERSION AF231021.1 GI:133182796
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 EXKYOTTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2494)
 AUTHORS Sharif, P. J., Kanai, N., Wang, L. Y., Vreeke, T. M. and Parker, C. J.
 TITLE Identification and characterization of a novel gene that is
 JOURNAL upregulated in leukemia cells by nitric oxide
 REFERENCE 2 (bases 1 to 2494)
 AUTHORS Sharif, P. J., Kanai, N., Wang, L. Y., Vreeke, T. M. and Parker, C. J.
 TITLE Direct Sub-Region
 JOURNAL Submitted (02-FEB-2000) Medicine, University of Utah and SLC VA
 Medical Centers, Box 151M, 500 Pothill Boulevard, Salt Lake City,
 UT 84143, USA

FEATURES
 Location/Qualifiers
 1..2494
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 1309..2192
 /note="Similar to the Homo sapiens ribonuclease/angiogenin
 inhibitor; expressed in hematopoietic cells; upregulated
 by nitric oxide in leukemia cells"
 /codon_start=1
 /product="leucine-rich-repeat protein RNO2"
 /protein_id="AAK4942.1"
 /db_xref="GI:133182796"
 /translation="MSCAMWHTSVSPATQAKAGLLPRLRLWKRCRISASACD
 LSAALANKLITMDLGSNGVFGPMVLCGLRHPDCLQICLRKCOLLESACGDEM
 ASVGLNPHVLELDLGNALDGLLCCGLRHPVCLRLTALKICRUTRAACDSEA
 STSVNGLRELDLGLNGLGLRHPVCLRLTALKICRUTRAACDSEA
 TGLNGLRHPVLELDLGNALDGLLCCGLRHPVCLRLTALKICRUTRAACDSEA
 RVTNPHYDGLG"

BASE COUNT 634 a 636 c 699 g 525 t
 ORIGIN

Query Match 15.9%; Score 783.6; DB 9; Length 2494;
 Best Local Similarity 90.0%; Pred. No. 4e-70;
 Matches 884; Conservative 0; Mis-matches 19; Indels 79; Gaps 1;
 QY 2495 ACCAG 2554
 Db 2042 ACCAG 1093
 QY 2555 CATTCACAGATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGT 2614

[illegible]

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
Rattus.
1 (bases 1 to 223344)
Muzny D, Warde E, Metzker M, Lee S, Abramson S, Adams C, Alder C, Allen C, Allen B, Alstrooms S, Amin A, Anguiano D, Anyalebechi V, Ayagari A, Aydogdu M, Baca S, Baden H, Baldwin C, Bandaru S, Barber M, Barstead M, Behrmann P, Bisswa S, Blair C, Blaser K, Blomberg K, Blyth P, Brown M, Bryant S, Burch C, Burch J, Burrell K, Calderon E, Chandanas V, Carter K, Cavasotto C, Ceasar H, Center A, Chacko M, Chavez D, Chen K, Chen R, Chen Y, Chen Z, Chu J, Clarke J, Coker S, Cox M, Coville M, Cree A, C'Struzo A, Davila M, Davis C, Day Carlsil J, De Anda C, Dederich D, Delgado O, Denson S, Deramo C, Ding Y, Dinn H, Divya K, Draper H, Egan-Pacha S, Dunn A, Durbin K, Duval B, Evans K, Egan A, Eskott M, Eugene C, Evans C, Falls C, Fan G, Ferrandez S, Fibley M, Flagg N, Forbes D, Foster M, Foster P, Fraser C, Gabis A, Gama R, Garcia A, Garner T, Garza M, Gebregeorgis E, Geer K, Gill R, Grady M, Guerra W, Guvarra W, Guravich P, Haaland W, Hae C, Hamilton C, Hamilton K, Haray Y, Havak P, Hawes A, Henderson N, Hernandez C, Hernandez A, Hires S, Hladun S, Hodgson A, Hoques M, Hollis B, Howell S, Hsly S, Hume J, Idelberg D, Jackson A, Jackson L, Jacob L, Jiang H, Johnson B, Johnson R, Colivet A, Karpach S, Kelly S, Kelly S, Khan Z, Kings J, Kovar C, Kowis C, Kratt C, L, Lepok H, Levay J, Lewis L, Li Z, Liu C, Liu J, Liu M, Liu Y, London P, Longacre S, Lopez J, Lorenzshwa L, Collapsed H, Lozados R, Lu X, Ma J, Maheshwari V, Mahindaratne M, Mahmood M, Malloy K, Manjun A, Vangun B, Mapua P, Martin K, Matt R, Martinez E, McInerney S, McLeod W, McNeil T, Meenen E, Milosavljevic A, Minei G, Mitja E, Montemayor J, Moore S, Morris M, Morris X, Morris S, Mridada M, Murphy M, Nait L, Navkevis C, Neal D, Newton N, Nguyen N, Norris S, Nwackelash O, Okwuno S, Ojannuragoon A, Pali S, Parks K, Pasternack S, Paul H, Perez A, Perez L, Frankson C, Pfaff P, Poindestet A, Popovich D, Primus E, Pu L, Puzos M, Qureshi A, Rachin E, Reeves K, Regier M, Reid R, Reilly B, Reilly M, Ren Y, Reuter X, Richards S, Riggs P, Rivers C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S, Sanders M, Savary G, Scherer S, Scott G, Shatsman S, Shen H, Shetty S, Shivasthayan A, Sisson I, Sitter C, Smajic D, Sneed A, Sodergren E, Song X, Sorelle R, Sosa J, Stamile M, Strong R, Sutton A, Syvack A, Taber P, Taylor C, Thomas N, Thomas S, Tingey A, Trejos Z, Usmani K, Valas P, Vera V, Villasana B, Waldron L, Walker H, Wang C, Wang S, Warren J, Warren R, Wei X, White F, Williams G, Wilson R, Wleczyk R, Wooden H, Worley K, Wright S, Wu J, Yakub S, Yen J, Yoon S, Yoon V, Yip P, Zhang J, Zhou X, Zhao S, Dunn D, von Kiedelshagen A, Weiss R, Smith D, Holt R, Smith H, Weinstock G, and Gibbs R. A.
Direct Submission
Unpublished
2 (bases 1 to 223344)
Worley K
Submitted 12-JAN-2002; Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 223344)
Rat Genome Sequencing Consortium.
Submitted 13-MAY-2003; Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced g1:23609300.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.tgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly. A 'contig-scaffold'. Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separate

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GJ02
Center clone name: CH210-22344
----- Summary Statistics
Consensus quality: 216992 bases at least Q40
Consensus quality: 214628 bases at least Q30
Consensus quality: 216538 bases at least Q20
Estimated insert size: 52520; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum of contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* See http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html.
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 1 contigs. Gaps between the contigs
  * are represented as runs of N. The order of the pieces
  * is believed to be correct as given, however the sizes
  * of the gaps between them are based on estimates that have
  * provided by the submitter.
  * This sequence will be replaced
  * by the finished sequence as soon as it is available and
  * the accession number will be preserved.
  * 223344: contig of 223344 bp in length.
  * Location/Qualifiers
    loc: 223344
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230 223820"
    990..1643
  misc_feature
    /note="clone boundary"
    clone_end:17
  site:EcoRI
    end sequence:PWBCD101023"
    77257..778553
  /notes="wgs contig"
  misc_feature
    219210..223047
    /note="clone boundary"
    clone_end:156
  site:EcoRI
    end sequence:PWBCD101023"
    225799..223344
  /notes="wgs end extension"
    clone_end:156"
  BASE COUNT  58709 a 51139 c 49375 g 54884 t 56567 others
  CIPHER

Query Match:      14.5%   Score 711.87   DB 27   Length 223344
Best Local Similarity  69.0%   Prod No. 580154
Matches 1138;   Conservative   0;   Mismatches 357;   Indels 208;   Gaps 2;

QY  1020  AGATCCCGAGAACCTACAGGAGCTATGTCGAGAGAAATCTGTCATGAGAGACCG 1079
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  150339  AGATCCCGAGAACCTACAGGAGCTATGTCGAGAGAAATCTGTCATGAGAGACCG 150398

QY  1580  CAATCCCGGCTAGGGAAATGTCACCTAGTCACCGGATATGTCGAGAGACCG 1139
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  150399  CAATCCCGGCTAGGGAAATGTCACCTAGTCACCGGATATGTCGAGAGACCG 150458

QY  1580  GAAGGACACTCAATCTAICTAGGACATATGTCGAGAGACCG 1159
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  150459  GAAGGACACTCAATCTAICTAGGACATATGTCGAGAGACCG 150518

```

```

QY  1200  TCCGAGACCTGTGGACACACAGGCTAGTCCCATTAAGATAGAGACCTCTTTGAGCCAGA 1259
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  150519  TCCGAGACCTGTGGACACACAGGCTAGTCCCATTAAGATAGAGACCTCTTTGAGCCAGA 150578

QY  1260  CAGGAGAGATGCGAGATGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1313
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  150519  TGAAGAACTGCGAGAGATGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150568

QY  1314  ----- 1313
DB  150639  GTGATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 150698

QY  1314  ----- ANAGAGATGAACACAGAGTCCGAGAGCTCTTCCAGAGCA 1345
DB  150639  GTTGGATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150758

QY  1346  CATTCAGAGATTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1405
DB  150759  TGTACATGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150818

QY  1406  CCGAGTTCGCGAGGCTCTCTTTCATCATCGAGGCTTGGATGAGTCAAGGCTTCTT 1465
DB  150819  CCGAGTTCGCGAGGCTCTCTTTCATCATCGAGGCTTGGATGAGTCAAGGCTTCTT 150878

QY  1486  CCGATGATCTTCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1525
DB  150879  CCGATGATCTTCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150938

QY  1526  TCTTAACAGCTTAATGGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1585
DB  150939  CTTTGGAGGCTGATTTGGAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150998

QY  1586  GCGCAGCTTTTGGAGAGCTTCCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1645
DB  150939  ACGCTGCTCACTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 151058

QY  1646  CCGTCTTCTGAGAGAGAGAGAGATATCTTCAAGATATTTCCAAATGCAAGAGCA 1705
DB  151059  CCGTCTTCTGAGAGAGAGAGAGATATCTTCAAGATATTTCCAAATGCAAGAGCA 151118

QY  1706  GCGCGCGCAGCTCTCATATCTGAGGAGACAGGACCTCTCTTCCAGCATGCTTCTCT 1765
DB  151119  ACCGAGTCAAGTCTTCTGCTTCACTGAGAGCTATGAGCGCTCTCTTACCCTGCTTCT 151178

QY  1766  CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1825
DB  151179  TCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 151238

QY  1826  GTTGAAGAGAGCTTCAAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1885
DB  151239  GTTGAAGAGAGCTTCAAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 151298

QY  1886  GAGCGGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1945
DB  151299  GAGCGGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151358

QY  1946  GTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2005
DB  151359  GTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151418

QY  2006  GCAAGGCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2065
DB  151419  ACATGGCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151478

QY  2066  CATTAATGCTGAG----- 2078
DB  151479  CATTAATGCTGAGAAATTTACAGCTTCTCCAGCTGAGTTTCCAGGAATTTCCACAGC 151538

QY  2079  ----- 2078
DB  151539  CATTAATGCTGAGAAATTTACAGCTTCTCCAGCTGAGTTTCCAGGAATTTCCACAGC 151598

```

```

CY 2079 -----AGGAGGTTCTGGCACTGATGAGGCGGATTCTGTTGTTGGACTCTGTAAGAGGGA 2131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DE 151593 CTGAGAAAGGAAGTCTTGGCCACACGCGACGCTTCTTCTGTTGGGCTCTCAATGAAGA 151658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 2132 GATCAGAGCCCACTGGAGAGAGGTCTCTTGGAGAGTCTGGGCGGCGGCGCATGAGATGGA 2191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DE 151659 GATGAGATGCTACCTCGAGAGGAATCTGCGCGGACCATCTTCCCTGAGGTGANGAGGA 151718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 2192 CTGTGTGCACTGGATCCAAAGCAAGCTTGAAGTGAAGGTTGATGCTGATGAGGAGGCTC 2251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DE 151719 AGGATGGCGTGGATCCAAACCAAGCTTGAAATGAGGATTTGATGCTGAGCATGCTC 151778
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 2252 CTGCGGACTTTTACGTTGCTGACGAGATTCGAGAGGATGAGTAAATATATATATAGGAGGCT 2311
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DE 151779 CCTGGAGCTCTCTAGCTGTTGTAAGATGCAAGAGTAAATATATATATATATATATATAT 151838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 2312 GAGCCACTTTCAGTGTATGTTGTGAGCAATATGCTGATGAGATGAGGATGATGCTCTC 2371
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DE 151839 GAGCCACTTTTCAGGTGTTGTTGTCAGAAAGCTCTTACGAGATGAGAGCATGCTGCTG 151898
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 2372 CTGCTGTCTGTAAGCGCTGCGAGAGGAGGAGTGTGCACTTGTATGGGCGCCACTA 2431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DE 151899 CTGCTGTTGTGCAAGCTACTGAGAGGCAATAGATGCTTACTTGTATGGGAGTCCCTA 151958
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 2432 CAGCGCGGAGGGGAGAGACCGCGGAGTGTGCTGATAGATGAGACAGCTGT 2484
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DE 151959 TATGACAGGGGCGAGAGATGGCTGGCGCAAAATCTTCAATGATGAGAGCAAT 152011
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: October 31, 2003, 09:13:14.
job time : 19581 secs

TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
MEDLINE
PUBMED
COMMENT
Contact: Mahairas GG, Wallace CC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: twallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Peter de Jong
(peter@dojng.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.html)
or from Research Genetics (<http://inforesgen.com>). BAC end web Server:
<http://www.hsc.washington.edu>
Plate: 825 row: N column: 10
Seq primer: SP6
Class: BAC ends
High quality sequence stops: 496
Location/Qualifiers
1..496
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate:825 Col:10 Row:N"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBAC3.6; Site: 1: EcoRI, Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI-MethIase. Size selected DNA was cloned into the
pBAC3.6 vector at EcoRI sites"
BASE COUNT 104 a 140 c 151 g 98 : 6 others
ORIGIN
Query Match 7.7%; Score 375.2; DB 28; Length 496;
Best Local Similarity 95.1%; Pred. No. 21;
Matches 387; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 2073 TGTGAGAGAGCTTCTGGCATACCAAGAGCTTGGTTTGGACTCTTATCGACAGAG 2132
DB 91 TGTGAAGAGAGTCTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 150
QY 2133 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2192
DB 151 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 210
QY 2193 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2252
DB 271 TGTGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 270
QY 2253 TGTGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2312
DB 330 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 390
QY 2312 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2372
DB 390 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 450
QY 2412 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2479
DB 497 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 497
RESULT 7
AQ726243

LOCUS
DEFINITION
Accession
Version
Keywords
Source
Organism
Reference
Authors
Title
Journal
Medline
Pubmed
Comment
Contact: Mahairas GG, Wallace CC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: twallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Peter de Jong
(peter@dojng.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.html)
or from Research Genetics (<http://inforesgen.com>). BAC end web Server:
<http://www.hsc.washington.edu>
Plate: 984 row: E column: 3
Seq primer: SP6
Class: BAC ends
High quality sequence stops: 496
Location/Qualifiers
1..496
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate:984 Col:E Row:E"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBAC3.6; Site: 1: EcoRI, Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI-MethIase. Size selected DNA was cloned into the
pBAC3.6 vector at EcoRI sites"
BASE COUNT 110 a 138 c 147 g 103 : 6 others
ORIGIN
Query Match 7.6%; Score 375.2; DB 28; Length 496;
Best Local Similarity 95.1%; Pred. No. 21;
Matches 386; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 2073 TGTGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2132
DB 91 TGTGAAGAGAGTCTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 150
QY 2133 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2192
DB 151 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 210
QY 2193 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2252
DB 271 TGTGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 270
QY 2253 TGTGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2312
DB 330 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 390
QY 2312 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2372
DB 390 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 450
QY 2412 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 2479
DB 497 AGCAGAGAGAGCTTCTGGCACTCCAGCAGCGCTTCTGTTTGGACTCTTAAACAGAG 497
RESULT 7
AQ726243

[illegible]

14 467 AAAGGGGGGGGGGAAAACCCGAGGGGTTTCAGAAAAGCCTCCACCCCTGTGGGC 526

Search completed: October 31, 2003, 06:57:55
Jcr time : 9474 secs

GenCore Version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: October 31, 2003, 03:57:44 : Search time 42 seconds
without alignment
263.125 Million cell updates/sec

Title: US-10-028-374-2

Interfer score: 2377

Sequence: 1 MNQATECSMQLPFCWPE.....GVKILVGLRHPNKKQLR 449

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1137863 seqs, 158724573 residues

Total number of hits satisfying chosen parameters: 1:07953

Minimum DB seq length: 3

Maximum LP seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A Geneset 10Jun01*

1:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1985.DAT*
2:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1981.DAT*
3:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1982.DAT*
4:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1983.DAT*
5:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1984.DAT*
6:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1985.DAT*
7:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1986.DAT*
8:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1987.DAT*
9:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1988.DAT*
10:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1989.DAT*
11:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1990.DAT*
12:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1991.DAT*
13:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1992.DAT*
14:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1993.DAT*
15:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1994.DAT*
16:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1995.DAT*
17:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1996.DAT*
18:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1997.DAT*
19:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1998.DAT*
20:	/SID1/gcgdata/geneseq/geneseqp-emb/AA1999.DAT*
21:	/SID1/gcgdata/geneseq/geneseqp-emb/AA2000.DAT*
22:	/SID1/gcgdata/geneseq/geneseqp-emb/AA2001.DAT*
23:	/SID1/gcgdata/geneseq/geneseqp-emb/AA2002.DAT*
24:	/SID1/gcgdata/geneseq/geneseqp-emb/AA2003.DAT*

Proj. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Match	Length	DB	ID	Description
1	2377	100.0	449	23	ABB77910	Amino acid sequenc
2	2344.5	98.6	1561	23	AA016590	Human PYRIN-3 prot
3	2334	98.2	1099	23	AA017857	Pyrin domain conta
4	1157.5	48.7	1034	22	AA027514	Human PYRIN-1 prot
5	1157.5	48.7	1034	24	AA028003	Human PYRIN-1 prote
6	1039.5	43.7	565	23	AB097475	Human nucleic acid
7	853	35.9	168	22	AB045470	Novel human diagno
8	816	34.3	920	23	AB045470	Human MDR-13 prot
9	510	34.1	519	23	AA017871	Pyrin domain conta

Human NB-ARC and C
Human NAC gamma of
Human CARD 7 polyp
Human leucine-rich
Leucine-rich repea
Human leucine-rich
Human caspase-1
Human caspase-1
Human NB-ARC and C
Human NAC beta 1sc
Human G-protein:cc
Human NB-ARC and C
Human NAC beta 1sc
Human caspase-1
Pyrin domain conta
Pyrin domain conta
Gene 28 Human seer
Human gene 214-enc
Human secreted pro
Human PYRIN-11 pro
Human leucine-rich
Leucine-rich repea
Arginyl-histidinyl-aso
Human NOL18a prote
Human G-protein:cc
Human ORF2121
Human nucleotide b
Human secreted pro
NOVX related prote
Pyrin domain conta
Human PYRIN-3 prot
Human activated T-
Pyrin domain conta
Human leucine-rich
Pyrin domain conta

ALIGNMENTS

RESULT 1

ABB77910
ID ABB77910 standard: Protein; 449 AA.

AC ABB77910;

XX CT CDT-2002 FIRST ENTRY.

XX DE Amino acid sequence of human leucine-rich repeat protein HLRPM.

XX KW Human; leucine-rich repeat; HLRPM; proliferative disorder;
XX KW immune condition; apoptosis; signal transduction; autoimmune disease;
XX KW haematopoietic cell disease; graft versus-host disease; allergy; asthma;
XX KW cardiovascular disorder; neurological disease; pheromone;
XX KW pulmonary disease; chronic obstructive pulmonary disease;
XX KW allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;
XX KW haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;
XX KW inflammatory disorder; systemic lupus erythematosus;
XX KW cardiovascular disease; cancer.

OS Homo sapiens.

XX PH Key location/Qualifiers
XX FT Domain 144..159 /note="transmembrane domain"

XX PK W020025201: A2.

XX PD 04-JUL-2002

XX PF 20-DEC-2001; 2001W0-US44740.

XX

22-DEC-2000; 2000US-257773P.
 (PRTM) BRISTOL-MYERS SQUIBB CO.
 Patel, C., Ramarathnam, C., Minicini, G.
 WPI: 2002-366676/60.
 N-PSDB; AB:59133.
 New HLRRAV1 nucleic acids for preventing, treating or ameliorating e.g.
 proliferative disorders, immune conditions, a disorder related to
 aberrant apoptosis modulation or developmental disorders
 Claim 20; Fig 1A-E; 37pp; English.
 The present sequence represents a human leucine-rich repeat containing
 protein, designated HLRRAV1. HLRRAV1 polypeptides and polynucleotides are
 useful for preventing, treating or ameliorating a medical condition such
 as a proliferative disorder, immune condition, or a disorder related to
 aberrant apoptosis modulation, either directly or indirectly, and in
 modulating signal transduction activity in various cells, tissue and
 organisms. They are also useful for treating, preventing, or diagnosing
 diseases of haematopoietic cells, autoimmune disease, graft-versus-host
 disease, allergic conditions (e.g. asthma), cardiovascular disorders,
 and neurological diseases, and for increasing the organism's ability to
 synthesize and/or release phenolamines. The polypeptide may also be used
 in treating, preventing or ameliorating pulmonary disease (e.g. chronic
 obstructive pulmonary disease, allergic rhinitis), or bronchial
 hyperresponsiveness; reproductive disease; haematopoietic disease;
 platelet disorders (e.g. Bernard Soulier syndrome, non-infectious,
 disorders (e.g. innate immunity to bacterial pathogens, or adaptive
 immune response), immune and inflammatory disorders (e.g. systemic lupus
 erythematosus), cardiovascular diseases and cancers. HLRRAV1 nucleic
 acids may further be used in chromosome identification or mapping, as a
 chromosome marker, as molecular weight markers, as diagnostic probes, in
 gene therapy, in raising anti-DNA antibodies, or as antigens for
 eliciting immune responses.

Db 42: LYRNALSPGVKLCQGRHHPNCKLQNL 449
 RESULT 2
 AAC15590
 ID AAC15590 standard, Protein: 1061 AA
 XX AC AAC15590:
 D7 31-OCT-2002 (first entry)
 XX DE Human PYRIN-8 Protein #2.
 XX KW Human; Gene therapy; PYRIN; stress-related response; apoptotic response;
 KW inflammatory response; inflammatory disorder; immune system disorder;
 KW Crohn's disease; multiple sclerosis; cancer; leukaemia;
 KW autoimmune disorder; arthritis; neurological disease;
 KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
 KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
 KW transcription profiling; PYRIN-8.
 OS Homo sapiens.
 XX KW2002061549-AA.
 PC Q8-ALM3-2302.
 XX 31 JAN-2002; 2002W0-USC2967.
 XX 31 JAN-2002; 2001US-245231P.
 PP 10-SEP-2002; 2001US-318445P.
 XX PA (MILL) MILLERIN PHARM INC.
 PA (AMHP) WYETH.
 XX Bertin J, Wang W, Blatchner M;
 WPI: 2002-627477/67.
 DR N-PSDB; AAL44363.
 XX New PYRIN polypeptides and nucleic acids useful for modulating and
 PT diagnosing stress-related, apoptotic and inflammatory responses, or for
 PT treating inflammatory and immune system disorders, cancers, or
 PT neurological diseases.
 XX Claim 8; Fig 6; 167pp; English.
 XX The invention comprises the amino acid and coding sequences of human
 CC PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
 CC useful for modulating and diagnosing stress-related, apoptotic and
 CC inflammatory responses. The PYRIN protein and DNA sequences are useful
 CC for treating inflammatory disorders and immune system disorders (e.g.
 CC Crohn's disease, reactive arthritis, multiple sclerosis, contact
 CC dermatitis, psoriasis, graft rejection, allergies, viral infections and
 CC bacterial infections). Cancer (e.g. leukaemia), autoimmune disorders
 CC (e.g. systemic lupus erythematosus and arthritis), and neurological
 CC diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
 CC protein and DNA sequences may also be used in screening assays, detection
 CC assays (e.g. chromosomal mapping, tissue typing or forensic biology),
 CC predictive medicine (e.g. diagnostic assays, clinical trials and
 CC pharmacogenomics) and transcription profiling. The present amino acid
 CC sequence represents a human PYRIN-8 protein.
 XX SQ Sequence 1061 AA;
 Query Match 98.6%; Score 2344.5; DB 23; Length 1061;
 Best Local Similarity 90.9%; Pred. No. 218-232;
 Matches 449; Conservative 3; Mismatches 0; Indels 45; Gaps 1;
 QY 1 MNSQATECSMQLIFSCWPEPSAPQELIRVERLJFIDGDELKPSFHDQPGWCLQW 60
 Db 255 MNSQATECSMQLIFSCWPEPSAPQELIRVERLJFIDGDELKPSFHDQPGWCLQW 144

```
QY 61 EKRRPTTELLNSLRKKQLPELSLTTTPPTALEKZCHMLLEHPRHVEILGSESAERKEYF 120
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 115 EKRRPTTELLNSLRKKQLPELSLTTTPPTALEKZCHMLLEHPRHVEILGSESAERKEYF 374
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 121 KYFHNAEGAGOVNVRNPEFTKCVFVLVWVYCTCQQQEGGGLQTSRTTAV 180
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 375 KYFHNAEGAGOVNVRNPEFTKCVFVLVWVYCTCQQQEGGGLQTSRTTAV 434
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 181 YMYLLSMOPKPGAPRLOFFPNQGLSLAAGGLMKNKILFECDLKKHGLDGEDVSAP 240
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 435 YMYLLSMOPKPGAPRLOFFPNQGLSLAAGGLMKNKILFECDLKKHGLDGEDVSAP 494
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 241 LKVNIFQKDNCG ..... 285 ERS 285
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 495 LKVNIFQKDNCGERYYSFTHLSFVFFFAAAYVVFETKGAHPDQCVTRLLTEVAFSRS 554
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 256 FIALTSRFFGLNLETRSHDESKLWKVSNKYLKGLWGLGSKAQSOGSTLQGSLEFF 315
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 355 FIALTSRFFGLNLETRSHDESKLWKVSNKYLKGLWGLGSKAQSOGSTLQGSLEFF 614
QY 316 SCLYETCEBEFTCOALSHPOVIVGNSIAEKHEPMVSSFCFKKRCSAOVHLGYGATYSADG 375
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 615 SCLYETCEBEFTCOALSHPOVIVGNSIAEKHEPMVSSFCFKKRCSAOVHLGYGATYSADG 674
QY 376 EPARCSAGARTLLVQCPERTYLDAYSEHIAALCTNPMLIELSLYENALGSRGVKLLC 435
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 675 EPARCSAGARTLLVQCPERTYLDAYSEHIAALCTNPMLIELSLYENALGSRGVKLLC 734
QY 436 QGLRHENCKLQNL 449
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 735 QGLRHENCKLQNL 749

RESULT 4
AA027857
1C AA027857 standard; Protein; 1039 AA.
XX
XX AA027857
XX
XX 20-AUG-2002 (first entry)
XX
XX Pyrin domain containing protein NALP3/EVER 88
XX
XX Pyrin domain; PYD domain; anti-inflammatory; anti-Alzheimer's;
XX anti-atherosclerotic; anti-proliferative; anti-inflammatory;
XX neuroprotective; anti-arthritic; anti-osteoporotic; anti-osteoporosis;
XX nephroprotective; osteoporotic; neuroprotective; anti-atherosclerotic;
XX inflammation; Alzheimer's disease; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
XX osteoarthritis; glomerulonephritis.
XX
XX Unidentified.
XX
XX W020224066-A2.
XX
XX 21 MAY 2002.
XX
XX 10-OCT-2001; 2001WO-EP12543.
XX
XX 15-NOV-2000; 2000DE-1056687.
XX
XX 30-NOV-2000; 2000DE-1059595.
XX
XX (APCT : APOTEC RES & DEV LTD.
XX
XX Tschopp J, Mattinen P;
XX
XX W01: 2002-427091/45.
XX
XX N-PSNB; ANL47129.
XX
XX New DNA encoding protein with pyrin domain useful for treating
XX diseases involving impaired signal transduction, particularly
XX inflammation, also proteins and antibodies
```

```
XX
PS
CC The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PYD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arthritis, osteoporosis, bacterial or viral infections particularly meningitis
CC and pneumonia, multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a protein of the
XX invention.
XX
SQ Sequence 1039 AA.
Query Match 98.14; Score 2334; DB 23; Length 1039;
Best Local Similarity 90.78; Pred No. 146 231;
Matches 449; Conservative 0; Mismatches 0; Indels 46; Gaps 2.
QY 1 XNQSATEGSMQDLISCPSPSAPLQELRVPERILFIIDGFDLKPSEHDPQFWLCLW 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 228 XNQSATEGSMQDLISCPSPSAPLQELRVPERILFIIDGFDLKPSEHDPQFWLCLW 287
QY 61 EKRRPTTELLNSLRKKQLPELSLTTTPPTALEKZCHMLLEHPRHVEILGSESAERKEYF 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 288 EKRRPTTELLNSLRKKQLPELSLTTTPPTALEKZCHMLLEHPRHVEILGSESAERKEYF 347
QY 121 KYFHNAEGAGOVNVRNPEFTKCVFVLVWVYCTCQQQEGGGLQTSRTTAV 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 345 KYFHNAEGAGOVNVRNPEFTKCVFVLVWVYCTCQQQEGGGLQTSRTTAV 403
QY 181 YMYLLSMOPKPGAPRLOFFPNQGLSLAAGGLMKNKILFECDLKKHGLDGEDVSAP 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 438 YMYLLSMOPKPGAPRLOFFPNQGLSLAAGGLMKNKILFECDLKKHGLDGEDVSAP 497
QY 241 LKVNIFQKDNCG ..... 285 ERS 285
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 495 LKVNIFQKDNCGERYYSFTHLSFVFFFAAAYVVFETKGAHPDQCVTRLLTEVAFSRS 557
QY 256 FIALTSRFFGLNLETRSHDESKLWKVSNKYLKGLWGLGSKAQSOGSTLQGSLEFF 315
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 529 FIALTSRFFGLNLETRSHDESKLWKVSNKYLKGLWGLGSKAQSOGSTLQGSLEFF 587
QY 316 SCLYETCEBEFTCOALSHPOVIVGNSIAEKHEPMVSSFCFKKRCSAOVHLGYGATYSADG 375
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 588 SCLYETCEBEFTCOALSHPOVIVGNSIAEKHEPMVSSFCFKKRCSAOVHLGYGATYSADG 647
QY 376 EPARCSAGARTLLVQCPERTYLDAYSEHIAALCTNPMLIELSLYENALGSRGVKLLC 434
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 649 EPARCSAGARTLLVQCPERTYLDAYSEHIAALCTNPMLIELSLYENALGSRGVKLLC 707
QY 435 QGLRHENCKLQNL 449
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 735 QGLRHENCKLQNL 749

RESULT 4
AA027514
1C AA027514 standard; Protein; 1014 AA.
XX
XX AA027514
XX
XX 01-NOV-2001 (first entry)
XX
XX Human PYIN 1 protein.
XX
XX Human nucleotide binding site; NBS-1; PYIN-1; apoptotic; cytostatic;
XX dermatological; vasoregulatory; melanoma; gene therapy; ophthalmological;
XX cancer; leukemias; carcinoma; arthritis; viral infection; allergy;
XX autoimmune disease; systemic lupus erythematosus; SLE; nocturnal asthma;
XX glomerulonephritis; neurological disorder; Alzheimer's disease; AIDS;
XX amyotrophic lateral sclerosis; Parkinson's disease; Huntington's disease;
XX aplastic anemia; leukemia; meningitis; liver disease; Crohn's disease;
```

XX insulin-dependent diabetes; multiple sclerosis; Grave's disease; HIV;
 XX human immunodeficiency virus; tuberculosis; lepromatous leprosy;

CS Homo sapiens.

XX Key Location/Qualifiers

FT Key 1..87

FT Domain /label= Pyrin_domain

FT Modified site 3..5

FT /label= Protein kinase C phosphorylation site

FT Modified site 44..46

FT /label= Protein kinase C phosphorylation site

FT Modified site 93..98

FT /label= N-myristoylation site

FT Modified site 112..113

FT /label= Casein kinase II phosphorylation site

FT Modified site 117..118

FT /label= Casein kinase II phosphorylation site

FT Modified site 164..167

FT /note= cAMP and cAMP dependent protein kinase phosphorylation site

FT Domain 224..233

FT /label= P-loop

FT /note= "Kinase 1a domain"

FT Modified site 226..268

FT /label= Protein kinase C phosphorylation site

FT Modified site 227..232

FT /label= N-myristoylation site

FT Domain 261..357

FT /label= NBS domain

FT Modified site 269..272

FT /label= Casein kinase II phosphorylation site

FT Modified site 290..293

FT /note= cAMP and cAMP dependent protein kinase phosphorylation site

FT Domain 290..306

FT /label= Walker B box

FT /note= "Kinase 2 domain"

FT Region 325..327

FT /note= "GSD cell attachment sequence"

FT Domain 344..355

FT /note= "Kinase 3a domain"

FT Modified site 347..349

FT /label= Protein kinase C phosphorylation site

FT Modified site 426..428

FT /label= Protein kinase C phosphorylation site

FT Modified site 431..435

FT /label= Protein kinase C phosphorylation site

FT Domain 446..449

FT /label= Dileucine motif

FT Modified site 491..496

FT /label= N-myristoylation site

FT Modified site 522..525

FT /label= Casein kinase II phosphorylation site

FT Domain 532..534

FT /label= Dileucine motif

FT Domain 559..560

FT /label= Dileucine motif

FT Modified site 588..591

FT /label= Casein kinase II phosphorylation site

FT Modified site 592..595

FT /note= cAMP and cAMP dependent protein kinase phosphorylation site

FT Modified site 595..597

FT /label= Protein kinase C phosphorylation site

FT Domain 606..607

FT /label= Dileucine motif

FT Region 618..626

FT /note= "peroxisomal targeting signal"

FT Modified site 624..627

FT /label= Casein kinase II phosphorylation site

FT Modified site 654..657

FT /label= N-glycosylation site

FT 656..658

FT /label= Protein kinase C phosphorylation site

FT 657..660

FT /label= Casein kinase II phosphorylation site

FT 670..672

FT /label= N-myristoylation site

FT 740..743

FT /label= Casein kinase II phosphorylation site

FT 749..767

FT /label= Leucine-rich repeat

FT 768..791

FT /label= NBS domain

FT 810..763

FT /label= Casein kinase II phosphorylation site

FT 829..794

FT /label= Leucine-rich repeat

FT 871..821

FT /label= Leucine-rich repeat

FT 815..816

FT /label= Dileucine motif

FT 816..837

FT /note= "Leucine zipper pattern"

FT 821..824

FT /label= Dileucine motif

FT 826..840

FT /label= Leucine-rich repeat

FT 854..876

FT /label= Leucine-rich repeat

FT 888..893

FT /label= N-myristoylation site

FT 891..908

FT /label= Leucine-rich repeat

FT 911..914

FT /label= N-glycosylation site

FT 923..924

FT /label= N-myristoylation site

FT 931..935

FT /label= Leucine-rich repeat

FT 921..924

FT /label= Casein kinase II phosphorylation site

FT 929..930

FT /label= Dileucine motif

FT 934..935

FT /label= Dileucine motif

FT 943..967

FT /label= Leucine-rich repeat

FT 951..953

FT /label= N-glycosylation site

FT 951..953

FT /label= Dileucine motif

FT 951..953

FT /label= Protein kinase C phosphorylation site

FT 968..991

FT /label= Leucine-rich repeat

FT 973..973

FT /note= cAMP and cAMP dependent protein kinase phosphorylation site

FT 977..998

FT /label= Dileucine motif

FT 1014..1017

FT /label= Casein kinase II phosphorylation site

FT 1018..1021

FT /label= N-myristoylation site

XX W020161005-A2.

PN 23-AUG-2001

XX 16-FEB 2001; 2001US 0541451

XX 17-FEB-2001; 2001US 0505673

PR 01-SEP-2000; 2000US 0643701

XX

XX MILLI MILLENNIUM PHARM INC.

XX Bertin J.

XX WP1: 2001-514773/56.

XX N-PSDB: AAD14323.

XX Novel polypeptide comprising NBS-1 or PYRIN-1 or a PYRIN-1 derivative of the pyrin domain protein family useful in screening and detection assays and for treating, e.g., cancer, viral infections, autoimmune disease, and Alzheimer's.

XX Claim 9: Fig 4: ilupp: English.

XX The invention relates to human NBS-1 nucleotide binding site; and PYRIN-1 protein and nucleic acids. NBS-1 and PYRIN-1 DNA, protein and their modulators are useful in the treatment of apoptotic and inflammatory disorders, cancer (leukemia, melanoma, carcinoma), viral infections (including herpesvirus and adenovirus), autoimmune diseases (systemic lupus erythematosus (SLE), immune-mediated glomerulonephritis, arthritis), neurological disorders (Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, Huntington's disease), aplastic anaemia, ischaemia, meningitis, liver diseases, NBS-1 and PYRIN-1 DNA, protein and their modulators are also used for the treatment of inflammatory and immune disorders such as chronic inflammatory diseases such as Crohn's disease, insulin-dependent diabetes, organ specific autoimmunity, including multiple sclerosis, Grave's disease, allergy, asthma, HIV, rubella, and leprosy. The present sequence is human PYRIN-1 protein.

XX Sequence 1034 AA;

Query Match 48.7%; Score 1157.5; CB 22; length 1034;

Best Local Similarity 45.3%; Pref. No. 9.66 112;

Matches 231; Conservative 93; Mismatches 117; Indels 69; Gaps 7;

RESULT 5

ABJ08503

ID ABJ08503 standard; Protein; 1034 AA.

XX AC ABJ08503.

XX 22-MAY-2003 (first entry)

DE Hura PYRIN-1 protein.

XX Human: nucleotide binding site; pyrin domain; NBS-1; PYRIN-1;

XX caspase-1; anti-inflammatory; apoptosis; ASC; NF-kB; nuclear factor KB, LRR; leucine rich repeat; inflammatory disorder; familial cold urticaria, arthritis; inflammatory bowel disease; Crohn's disease; asthma; allergy; ulcerative colitis; rheumatoid arthritis; Lyme disease; Grave's disease; insulin-dependent diabetes; multiple sclerosis; contact dermatitis; psoriasis; graft versus host disease; food allergy; conjunctivitis; chronic obstructive pulmonary disease; food allergy; conjunctivitis; chronic obstructive pulmonary disease; food allergy; conjunctivitis; helminthic infection; Leishmaniasis; viral infection; HIV infection; bacteria; infection; tuberculosis; leprosy; chromosome 1q44.

XX Homo sapiens.

XX US20-22187922-A1.

XX 12-DEC-2002.

XX 22-APR-2002; 2302US 0127516.

XX 17-FEB-2000; 2300US-056567.

XX 01-SEP-2000; 2300US-0653901.

XX 26-SEP-2001; 2301US 0964955.

XX 20-DEC-2001; 2301US 0327629.

XX (BERTIN); BERTIN J.

XX (MANU); MANU; G. A.

XX Bertin J. Manu GA;

XX WP1: 2001-514773/56.

XX N-PSDB: AAD14323.

XX Identification of compound that binds to polypeptide (for e.g. a PYRIN protein), useful for treating disorders associated with inappropriate apoptosis, for e.g. inflammatory disorder.

XX Claim 1: Fig 4: 74pp: EMB:158.

XX The invention relates to a new method for the identification of a compound that binds to a human PYRIN-1 (an apoptotic signalling molecule appearing as ABJ08503) comprising contacting the polypeptide or a cell expressing the polypeptide to a test compound and determining whether PYRIN-1 binds to the test compound, or determining the effect of the test compound on the activity of PYRIN-1. In a similar manner, compounds are isolated which modulate the binding of PYRIN-1 to ASC (comprising testing the compounds against the pyrin binding domains of both PYRIN-1 and ASC (not defined)), modulate the activity of NF-kB (nuclear factor KB), and which modulate the ASC-mediated activation of NF-kB (comprising: (a) measuring the binding of a test compound to the LRR (leucine rich repeat) domain of PYRIN-1; and (b) measuring the activation of NF-kB in a cell expressing ASC and PYRIN-1 in the presence and absence of the LRR domain binding compound). The candidate modulators may be useful for treating an inflammatory disorder. Also included is a method (M6) for identifying a modulator of caspase-1 activity. The identified compound can be used to modulate the polypeptide's activity (ASC and NF-kB activities in a patient). By modulating the expression or activity of the polypeptide (PYRIN-1 and ASC), a disorder associated with inappropriate apoptosis (inflammatory disorders e.g. familial cold urticaria, arthritis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, rheumatoid arthritis, Lyme disease, insulin-dependent diabetes, multiple sclerosis, Grave's disease, contact dermatitis, psoriasis, graft versus host disease, asthma, chronic obstructive pulmonary disease, allergies (e.g. food

PA

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

allergies), conjunctivitis, helminthic infection (e.g. Leishmaniasis), viral infections such as HIV infection, and bacterial infections such as tuberculosis and leprosy) can be treated. Human NBS-1 protein (nucleotide binding site protein 1) contains a pyrin domain which was used to identify genomic sequences containing the PYRIN-1 gene. The gene for pyrin-1 is located on chromosome 1q44. The present sequence represents human PYRIN-1.

XX
SQ Sequence 1034 AA;

Query Match 48.7%; Score 1157.9; DB 24; Length 1034;
Best Local Similarity 45.1%; Pred. No. 9.4e 110;
Matches 211; Conservative 93; Mismatches 117; Indels 69; Gaps 7;

QY 6 TRSNCGLIFSCWPSASLQELIRVPELIIIDFENLKRSHPCPGWVLCWEKRP 65
DB 266 TQSGGLINSCCPDPPIKIVKPSITFMDGDFELQGFDEHIGPCTDCKAER 325
QY 66 TELLINSLIRKIKLPELSLITTRTAJELKRLHLEHPPHVLGFSAEKREYKYPH 125
DB 324 GDLSLIRKIKLPEASLITTRPALEKLOKHLCHPEHVLGFSAEKREYKYPH 385
QY 126 NAECAQOVYVDRNPLETNTFVPLVGVVCTCLQOLEGGLIRQSRITTAVMYL 185
DB 386 DEACAAAEFLIQENVEFTMGFTPLVGMVGTGLKQVESKSLAQSKITAVYVPEL 445
QY 186 LSLMOPKGPAPLQPPNPGRLCSLAADGLMKNKILFEHCDLRKSLDGVSAFKNVI 245
DB 446 SLLQFGSGQEHGICAHLWGLCSLAADGTWVKILFEESDIPNHCQKADVSATRNVL 505
QY 246 FQKDNCR--SFLAT----- 260
DB 506 FQKVEPCFYPFIHMTFQFFFAFMYVLEEEKKPTNVEGSKLPKSDATVLELVGK 565
QY 261 -----SRFPGLINSETHLERSLQWVSHIKULLQWCSMAQSSLSLQGS 311
DB 566 FEGYLIPTVRFGLGVNFRSYLENKLSLCSQGLRLKLLKLVKAKAKKGLQPSQ 625
QY 312 DEFFSCVETQEETFCQALSHFQVTVSHAKMKNVSSHTKPKCSAGVHLVGATY 371
DB 626 DELFYCLYWCEDFVCRANDYFKLETNLSPTPDHMYSSPTENCHVETLSL-GTGH 683
QY 372 SAGEGRARCSAGPILVQ-LPESTVLLDAY--SEHIAALT-----TNPKEIF 418
DB 684 NYPESEEEKEGRHLDVQCVLPSSHAAGHLVKNHLSSTSFNGLFVSLSTQSGLTE 743
QY 419 LSTVRAALSGRWKLLCGGLEHINQWLN 448
DB 744 LLSLSASLGDGPKRVLCETGICHPKNIPE 777

RESULT 6
AB097475 ID AB097475 standard; Protein: 565 AA.
AC AB097475;
XX
XT 16 DEC-2002 (first entry)
XX
DE Human nucleic acid associated protein, NAAP, from INCVTE 10.429930CD1.
XX
KW Human; nucleic acid associated protein; NAAP; cancer;
KW cell proliferative disease; cancer; atherosclerosis; hepatitis;
KW neurological disorder; Parkinson's disease; Alzheimer's disease;
KW stroke; epilepsy; developmental disorder; renal tubular acidosis;
KW anaemia; glaucoma; hypothyroidism; autoimmune disorder; AIDS;
KW inflammatory disorder; acquired immunodeficiency syndrome; allergy;
KW atopic dermatitis; arthritis; bacterial infection; viral infection;
KW parasitic infection; protozoal infection; fungal infection.
XX
CS Homo sapiens.
XX
XX W0200272630-A2.
EN

XX 19-SEP-2002.
XX 07-FEB-2002; 2002W0-USC3844.
XX 09-FEB-2001; 2001US 2691-8P.
XX 21-FEB-2001; 2001US 2739-3P.
XX 22-FEB-2001; 2001US-2708-8P.
XX 23-FEB-2001; 2001US-2711-9P.
XX 27-MAR-2001; 2001US 2740-1P.
XX 12-APR-2001; 2001US 2834-6P.
XX 09-NOV-2001; 2001US 3446-5P.
XX (INCVTE GENOMICS INC.
XX Thornton M, Hafalia Aa, Lu DM, Arvizu C, Swarnakar A, Lu Y;
XX Warren BA, Baughn MR, Tang Y, Lee EA, Yao MG, Ramkumar J;
XX Khan FA, Gandhi AS, Ding L, Yue H, Gietzen KJ, Walla NK;
XX Thangavelu K, Elliott VS, Marquis JP;
XX WPI: 2002 723320/78.
XX N-PSDB: AB59719.
XX New human nucleic acid associated proteins (NAAP), useful for
XX diagnosing, treating and preventing diseases or conditions associated
XX with the aberrant NAAP expression, e.g. cancer, hepatitis, AIDS,
XX atherosclerosis, infections.
XX
XX Claim 1: Page 147-149; 149pp; English.
XX The invention relates to an isolated polypeptide comprising one of 10
XX human nucleic acid associated protein (NAAP-10), or a biologically
XX active or immunogenic fragment of the polypeptide, and their encoding
XX nucleic acid. Also included are a recombinant polynucleotide comprising a
XX promoter sequence operably linked to the polynucleotide, a cell
XX transformed with the recombinant polynucleotide, a transgenic organism
XX comprising the recombinant polynucleotide, an anti-NAAP antibody,
XX screening for a compound that is effective as an ant/agonist or
XX modulator of NAAP, generating an expression profile of a sample
XX containing the polynucleotides and an array comprising different
XX nucleotide molecules affixed on a solid substrate, nucleotide molecule
XX comprises a first oligonucleotide or polynucleotide sequence specifically
XX hybridisable with at least 10 contiguous nucleotides of the target (NAAP)
XX polynucleotide. The polypeptides and polynucleotides are useful in
XX diagnosing, treating and preventing diseases or conditions associated
XX with the decreased expression or overexpression of NAAP, such as cell
XX proliferative diseases (e.g. cancer, atherosclerosis, hepatitis),
XX neurological disorders (Parkinson's disease, Alzheimer's disease,
XX stroke, epilepsy), developmental disorders (renal tubular acidosis,
XX anaemia, glaucoma, hypothyroidism), autoimmune/inflammatory disorders
XX (AIDS (acquired immunodeficiency syndrome), allergies, atopic dermatitis,
XX arthritis and infections (e.g. bacterial, viral, parasitic, protozoal,
XX fungal) and many other diseases and disorders listed in the
XX specification. These are also useful in assessing the effects of
XX exogenous compounds on the expression of nucleic acid and amino acid
XX sequences of NAAP. The NAAP or its fragments are useful in screening
XX compounds for effectiveness as agonist or antagonist of the polypeptides,
XX or in altering the expression of the target polynucleotide and compounds
XX that specifically bind to or modulate the activity of the polypeptide.
XX The microarray is useful in monitoring or measuring protein-protein
XX interactions, drug target interactions, and gene expression profiles.
XX The present sequence represents an NAAP protein.

XX Sequence 565 AA;

Query Match 43.7%; Score 1039.5; DB 23; Length 565;
Best Local Similarity 82.1%; Pred. No. 5.9e-98;
Matches 207; Conservative 1; Mismatches 30; Indels 45; Gaps 11

QY 243 MN:FGKQINC
DB : MN:FGKQINC:PEPYYS:HLISFQFFFAFVYILHDEGGAGPDDVTRLLTEVAFSHSEFL 60

QY 258 ALTSRFLGGLNEETRSHEKSLKWKVSPH:KMOALLQWIOQAQSDGSLQGSLEFFSC 317
 DB 61 ALTSRFLGGLNEETRSHEKSLKWKVSPH:KMOALLQWIOQAQSDGSLQGSLEFFSC 120
 QY 318 LYEIOEEEFIOQALSHFQVIVYSNIASKVHEHVMVSFFLKPCEAQQVLYGATYSADGEC 377
 DB 121 LYEIOEEEFIOQALSHFQVIVYSNIASKVHEHVMVSFFLKPCEAQQVLYGATYSADGEC 180
 QY 378 RARCSAGANTLVOLPERTVLDAYSEHLAAALCTNPNIIELSLYR 437
 DB 181 RARCSAGANTLVOLPERTVLDAYSEHLAAALCTNPNIIELSLYR 240
 QY 438 LRHPNCKLQNER 449
 DB 241 LRHPNCKLQNER 252
 RESULT 7
 ID ABG54570 standard; Protein: 168 AA.
 AC ABG54570;
 DT 13-FEB-2002 (first entry;
 XX Move: human diagnostic protein #4561.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 CS Homo sapiens.
 XX
 PN W:200175067-A2.
 FC 11-OCT-2001.
 XF 30 MAR 2001; 2001WC-USC4631.
 XX
 PE 11-MAR 2000; 2000JS-US40217.
 PR 23 AUG-2000; 2000JS-5649167.
 XX
 RA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AA568757.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID No 34923; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotides (i) and
 CC polypeptide (ii) sequences. (i) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligonucleotides, and for chromosome
 CC and gene mapping, and in recombinant production of (ii). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (ii) is useful in gene therapy techniques
 CC to restore normal activity of (i) or to treat disease states involving
 CC (iii). (iii) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (ii) and its binding partners are useful in medical
 CC imaging of sites expressing (ii). (i) and (ii) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 168 AA;
 Query Match 35.9%; Score 853; DB 22; Length 168;
 Best Local Similarity 59.4%; Pred. No. 1.7e-79;
 Matches 167; Conservative 0; Mismatches 1; Indels 0; Gaps 0.
 QY 256 FLALTSRFLGGLNEETRSHEKSLKWKVSPH:KMOALLQWIOQAQSDGSLQGSLEFF 315
 DB 1 FLALTSRFLGGLNEETRSHEKSLKWKVSPH:KMOALLQWIOQAQSDGSLQGSLEFF 60
 QY 316 SCUYETQSEEFIOQALSHFQVIVYSNIASKVHEHVMVSFFLKPCEAQQVLYGATYSADG 375
 DB 6 SCUYETQSEEFIOQALSHFQVIVYSNIASKVHEHVMVSFFLKPCEAQQVLYGATYSADG 120
 QY 376 EDRARCSAGANTLVOLPERTVLDAYSEHLAAALCTNPNIIELSLYR 423
 DB 21 EDRARCSAGANTLVOLPERTVLDAYSEHLAAALCTNPNIIELSLYR 168
 RESULT 8
 ID ABG53254 standard; Protein: 920 AA
 AC ABG53254;
 DT 13-NOV-2002 (first entry;
 XX Human MDDT-13 protein SEQ ID NO:13.
 XX
 KW Human; MDDT; molecule for disease detection and treatment; anti-HIV;
 KW anti-arteriosclerotic; cytosstatic; neuroprotective; anti-parkinsonian;
 KW hepatotropic; laxative; cerebroprotective; antiinflammatory; virucide;
 KW antibacterial; fungicide; prozoocide; gene therapy; arteriosclerosis;
 KW cirrhosis; cancer; stroke; Alzheimer's disease; Parkinson's disease;
 KW Crohn's disease; constipation; acquired immunodeficiency syndrome; AIDS;
 KW infection.
 XX
 OS Homo sapiens.
 XX
 PN WC200264792-A2.
 XX
 PD 22-AUG-2002
 PF 04-JAN-2002; 2002WC-US00254.
 XX
 PR 05-JAN-2001; 2001US-260168P.
 PR 19-JAN-2001; 2001US-262366P.
 PR 19-JAN-2001; 2001US-262857P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Baughn MR, Warren BA, Henchell CD, Xu Y, Wallia NK, Ramkumar J;
 PI Rao MG, Lu Y, Yue H, Thangaveeru K, Tang TY, Qing L, Borowsky ML;
 PI Hafalila ACA, Lu DAW, Azimzai Y, Tran B, Nguyen DB, Burford N;
 PI Ison GH, Gururajan R, Gandhi AR, Elliott VS, Tran UK;
 XX
 DR WPI: 2002-674890/72.
 DR N-PSDB; ABQ75811.
 XX
 PT New purified polypeptides, molecules for disease detection and
 PT treatment (MDDT), useful for diagnosing or treating a disease or
 PT condition associated with decreased expression or overexpression of
 PT functional MDDT e.g. stroke.
 XX
 PS Claim 1; Page 128-130; 162pp; English.
 XX
 CC ABQ75799 to ABQ75814 encode the human molecules for disease detection
 CC and treatment 1 to 26 (MDDT-1 to MDDT-26) given in ABP53242 to ABP53267.
 CC MDDT's have anti-HIV, antiarteriosclerotic, cytosstatic, neuroprotective,

CC antiparkinsonian, hepatotropic, laxative, neuroprotective, virucide,
CC anti-inflammatory, antibacterial, fungicide and protease activities,
CC and can be used in gene therapy. MOR sequences can be used in the
CC diagnosis and treatment of a disease or condition associated with
CC decreased expression or overexpression of functional MOR e.g. cirrhosis,
CC arteriosclerosis, cancer, stroke, Alzheimer's disease, Parkinson's
CC disease, Crohn's disease, constipation, acquired immunodeficiency
CC syndrome (AIDS), or bacterial, viral, fungal or protozoal infections.
XX
XX Sequence 920 AA;

Query Match	34.31	Score 816	25 23	Length 923
Rest Local Similarity	38.58	Pred. No. 1.5e 74		
Matches 192	Conservative	76	Mismatches 159	Index's 72
				Gaps 10

```

7  ECSCMDLFCSCWPSAPJQBLIRVBERLLFTIDGDEKDEKPSHHFQSPWCLWEEKPT 66
DB 285  ERSAQAQJLKKOMPSTEGIEBIKQYQSSLLFIIDSEDEENFAFEERFALZEDTQHPV 314
QY 67  ELLNLSLRKKLGLPELSLITITPTALEKILRLEKLEHPRHVEILGFESEAKRKYVYKPHN 126
DB 315  SPKXSSLLRKQVLEASLIVTITTSKRLFOLEKSHYVEILGSHREBYVQPFED 374
QY 127  ABQACQVFNYYKONEPIFTKCFEPLKQVWNTYKIOGLEGORSELICTSKTTAVYMYELL 186
DB 318  KRWAKYFESSLSKSNEMFLSCQVFLVWAACTLEKQWBEKGDVITTCCTTALFTYVIS 434
QY 147  SLMGP-KXGAPLQPPNO--EGLCSLAALRLMNCLELFEQDLEPHIGLEGEDVSAPLN 242
DB 435  SLTFEYVQGSFSL---PQACLERLQVAAKQNTVYVYFRENLRRLCJGCSVSSPND 491
QY 243  MNILQKGI--NCF----- --RSPLALTER-- 262
DB 492  SNILCKQMEYCNVFTHLVQVCFMAASTYKMAKSMEARNPSCQPFEDKLSLQSTSYK 551
QY 263  -----FJGLKLETRSHLEKSLCKWKYSPIKMDLEQMLQSKAQDGSFTDQGSJEF 314
DB 552  MLTQMKCFJGLLNEBRYVQLETFNCKYMSLEKLEKLEQMEVLGNSTYSPSLQPLEL 612
QY 315  FSCVYETQFBEFIOCALSHFQVLIWENTASKVBNKYSRLEKQKQSAQVYHNGATYSAD 374
DB 612  FHCYVFOKAFISQAMRCFQKVALNICEKILHVLVSSPFEKHCRLQPTIEL--SVTW 667
QY 315  GEDRAKSAGANTLIVQ:PERT-----VLLMAYLEHIAALITNPKLELSYENALQSR 429
DB 668  FEKK-- --LKTSEJFNTWEEYVGLSLEQJLSALITKPLKPLKTKTLOVE 719
QY 410  CVKLLCOOSLRHFNCKLQNL 448
DB 720  G:VKLYKVLKSPKCKLOVL 738

```

RASUR: 9	
AAQ:7871	
ID	standard: Protein: 5:8 AA.

XX	
XX	AAG19841;
XX	AC
XX	DT
XX	26-AUG-2002 (first entry);
XX	
DE	Pyrim domain containing protein NAF13/PYD9 mm

XX	Pyruv domain; pyv domain; antinuclear; antiparkinsonian;
XX	antiartherosclerotic; antiparasitic; antibacterial; virucide;
KW	neuroprotective; antiarthritic; antithrombotic; antiaslatic;
KW	reprotoprotic; osteopathic; nociceptive; intracellular signal transduction;
KW	inflammation; Alzheimer's disease; infectious; psoriasis; asthma;
KW	antileukoproliferative; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
XX	osteoarthritis; glomerulonephritis;
XX	
CS	Unidentified.
XX	
PN	WC2002040609-A2.

WC 2240669-A2.

XX	23-MAY-2002.	
PP		
XX	30-OCT-2001: 2001WC EP:2545.	
XX		
PP	15-NOV-2003: 2003DE 1056687.	
PR	30-NOV-2003: 2003DE 1059595.	
XX		
XX	(APCT-): APOTECHE RES & DEV LTD.	
PA		
XX		
XX	Tschopp C, Martinez F;	
P1		
XX	WP: 2002-427833/45.	
XX	N-PSDB; AALAT143.	
CR		
XX		
XX	New DNA encoding protein with pyrin domain, useful for treating	
PT	diseases involving impaired signal transduction, particularly	
FF	inflammation, also proteins and antibodies .	
XX		
XX	Claim 5: Fig 1; 116pp; German.	
PS		
XX		
XX	The present invention relates to DNA and their encoded proteins, where	
CC	the proteins contain at least one PYD (pyrin) domain. These can be used	
CC	to treat diseases associated with impaired intracellular signal	
CC	transduction, particularly inflammation such as psoriasis,	
CC	arteriosclerosis, bacterial or viral infections (particularly meningitis	
CC	and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,	
CC	sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's	
CC	and Parkinson's diseases. The present sequence is a protein of the	
CC	invention.	

Sequence 5: 9 AA;

Query Match 34.1%; Score 210; DB 23; length 518;
Best Local Similarity 60.6%; Pred. No. 2.6e-74;

Matches	Conservative	42	Mismatches	57	Indels	0	Gaps	0
QY	6	TEUNQUILFSCMPERSAPQQLIRKVFRELLTIIIGFDELKPSHDPGGPWCJGWEKRP	65					
DB	264	TPESLAIVISCMPPAPPPCKIRKRFESKILFLNGSGFDELQCAFDEHIGEVCTQCKAVR	322					
QY	56	TE.....NS::TKKILPELS::TTTPALSKLHKLHPRHVEILGFSABERKVFYKYPH	125					
DB	324	QDILSSVTKKLEPKKAS::TTTPVALEKGLHLLDHPRHVSIIGFSABERKVFYKYPH	393					
QY	156	NAFAGGVENVVPEENPPLPFWCFVFLVWVNTCLQQQLGGGLLRQTSRTTAVVLYL	185					
DB	384	NELKAREAFRIIQENEVFWMCIFPLVCH::VTGKQWETGKSLAQKSTTTAVVFFL	443					
QY	186	LSLWPKFAGAPRQPPNCRGKCSLAADSLWQKLFEBEQLRKHGLGDECVSAFLNNKI	245					
DB	444	SSLLOSPPGTEHLEFSEYICGKCSLAADGKWKQKLFEBECLRKHGLQKTVSAFLRMV	504					

RESULT 10

AA72670
ID AA72670 standard; Protein; 1397 AA.

XX
AC

31-MAY-2001 (first entry)

XX
DE Human NB-AR^c and CARG containing protein (NAC) gamma isoform.

Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC; caspase-associated recruitment domain; CARD; TIM-Barrel-like domain; cysteine aspartic protease; apoptosis; cytokine production; cytokine receptor signaling; therapy; inflammatory disorder; sepsis; fibrosis; arthritis; cancer; adenocarcinoma; leukaemia.

CC processes, are principal effectors of apoptosis. CARD-5 and CARD-7 are used
 CC for screening modulators that modulate apoptosis, cytokine production,
 CC cytokine receptor signaling and other cellular processes. They can act
 CC as an immunogen for the production of polyclonal and monoclonal
 CC antibodies. They can also be used to diagnose and treat inflammatory
 CC disorders such as sepsis, fibrosis and arthritis and cancer pathologies
 CC such as adenocarcinomas and leukemias.

SV Sequence 1429 AA;

Query Match 31.9%; Score 759; DB 22; Length 1429;
 Best Local Similarity 37.6%; Pred. No. 2,3e-68;
 Matches 193; Conservative 71; Mismatches 165; Indels 69; Gaps 9;

QY 9 SMOGLFSCWPEAPLOELRVVERLFILOGLDELEKESRHHGPGKCLWELKRPTEL 68
 DB 379 SLAEELGKDTATPAPIRIQLSRPRLFIIRVDEPGWVLQEPSSSELCLHWSQPADA 438
 QY 69 LLSNLRKLLPELSLILITRTALEKLRHLEHRRHVEILGFSAEKKEVYKYNNAE 128
 DB 439 LLSLSLCKLILPEASEFLITARTALQNLPSQEQARWVEVLGFSSESRKEYFYRTDER 498
 QY 129 CAGQVKNVVRNEPFTNCVPLVGVVTCVCCGQKRAHMLPQTSRTTAVVYMLLSL 188
 DB 499 QAIRAFVYKSKKELWALCVVWVWENLACTLMQONKKEKLTITSLTCLLYLAQA 559
 QY 189 MQPKPGARLOPPNQRGLCSAAGLWKKILFEEDLERNGLGDEVSAPLNNKVFQ 247
 DB 559 LCAQPLGROL -----RLCSLAAGLWKKLTLSFHLERHSLDGLITFLNGILOE 612
 QY 248 -----KQIND -----ERS -----PLATSP 262
 DB 613 HPPLVSYPFHLCPEFFPAANSVLEDEKGRKSNVCTILEKTEAVSGHLEASTR 672
 QY 263 KFTMLAKETSHLEKSKWVSPKVKCLLWQKQKQKQKQKQKQKQKQKQKQKQK 322
 DB 673 PLGLSGSGRGKVENIFRPLDQ GPNNWCVTSQ ----LPHSELSNCTVETR 726
 QY 323 SEEFIOGLSHPOVAVWNTAKSEHMYVSRNG KPTFSQVVEHVMATYSADGERAKCS 382
 DB 727 KNTLQVMAHEEFGMC VETQSHLACTLTKPSHIVKLCOL -----EGRQHS 779
 QY 383 AGAHTLLVQLPERTVLDAYSHHAAALTNFNLHEDTFNLAISRGVKLLCGQLRHN 442
 DB 778 WASTPMV LFWVVEVDVAYKGLFSLSVI KVTGNI KETFLGKSTSHRAVKSUCKTARIR 814
 QY 443 CKLQNLK 449
 DB 817 CULSTLR 843

RESULT 12
 AAB2571
 ID AAB2571 standard; Protein; 1429 AA.
 XX AAB2571;
 AC AAB2571;

XX 23 JUL 2001 (first entry)
 DT Human CARD-7 polypeptide.

XX CARD-7; CARD-8; CARD-5; caspase recruitment domain; tumor; human;
 KW autoimmune disorder; antiinflammatory; immunosuppressive; anti-leish;
 KW antibacterial; antiviral; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Domain 329..645
 FT /notes "nucleotide-binding domain"
 FT Domain 333..341
 FT /notes "kinase 1A subdomain"
 FT Domain 424..413

FT /notes "kinase 2 subdomain"
 FT 454..463
 FT /notes "kinase 1a subdomain"
 FT 455..462
 FT /notes "motif 2 domain"
 FT 497..484
 FT /notes "leucine-rich domain"
 FT 834..863
 FT /notes "leucine-rich domain"
 FT 864..891
 FT /notes "leucine-rich domain"
 FT 893..911
 FT /notes "leucine-rich domain"
 FT 921..948
 FT /notes "leucine-rich domain"
 FT 950..976
 FT /notes "leucine-rich domain"

XX MO200130813.A1.

PN 03-MAY 2001.

XX 27 OCT-2002; 250260-US29746.

XX 27-OCT-1999; 90US 0428252.

XX INWILL : MILLENNIUM PHARM INC.

XX Bertin C;

XX WPI; 2001-34740/36.

XX N-PSOE; AAF83451.

XX Identifying a modulator of interaction between caspase recruitment
 domain CARD-7 and CARD-5, for treating autoimmune disorders,
 PT comprises treating the binding of CARD-7 and CARD-5 in the presence of
 the compound

XX Disclosure: Fig 1A-D; 83ff; Enj.165.

XX The invention relates to identifying compounds that modulate interaction
 CC between caspase recruitment domain (CARD-7 and CARD-5). The method
 CC involves measuring the binding of CARD-7 and CARD-5 in the presence of
 CC the compound (an increase in the binding of CARD-7 to CARD-5 in the
 CC presence of the compound compared to the binding in the absence of the
 CC compound indicates that the compound is a modulator of CARD-7-CARD-5
 CC interaction). Modulators of CARD-7 and CARD-5 expression or activity can
 CC be used to treat or diagnose disorders such as cancers, bacterial or
 CC viral infections, autoimmune disorders (systemic lupus erythematosus,
 CC immun-mediated glomerulonephritis or arthritis), inflammatory disorders,
 CC organ-specific autoimmune including multiple sclerosis, Hashimoto's
 CC thyroiditis, or Grave's disease, psoriasis, graft rejection, allergies.
 CC CARD-7 and CARD-8 are useful as modulating agents in regulating a variety
 CC of cellular processes including cell growth and cell death. The present
 CC sequence represents the human CARD-7, an intracellular protein.

XX Sequence 1429 AA;

Query Match 31.9%; Score 759; DB 22; Length 1429;

Best Local Similarity 37.6%; Pred. No. 2,3e-68;

Matches 193; Conservative 71; Mismatches 165; Indels 69; Gaps 9;

QY 9 SMOGLFSCWPEAPLOELRVVERLFILOGLDELEKESRHHGPGKCLWELKRPTEL 68

DB 379 SLAEELGKDTATPAPIRIQLSRPRLFIIRVDEPGWVLQEPSSSELCLHWSQPADA 438

QY 69 LLSNLRKLLPELSLILITRTALEKLRHLEHRRHVEILGFSAEKKEVYKYNNAE 128

DB 439 LLSLSLCKLILPEASEFLITARTALQNLPSQEQARWVEVLGFSSESRKEYFYRTDER 498

QY 129 CAGQVKNVVRNEPFTNCVPLVGVVTCVCCGQKRAHMLPQTSRTTAVVYMLLSL 186

DB 499 QAIRAFVYKSKKELWALCVVWVWENLACTLMQONKKEKLTITSLTCLLYLAQA 559

CC treating, or ameliorating a neural disorder or a disorder related to
 CC aberrant apoptosis modulation (either directly or indirectly), renal
 CC disorder, immune disorder (e.g. arthritis, asthma, acquired
 CC immunodeficiency syndrome (AIDS), or rheumatoid arthritis),
 CC haematopoietic, metabolic, reproductive, pulmonary or cardiovascular
 CC diseases, hyperproliferative disorders, neurodegenerative diseases
 CC (e.g. Alzheimer's disease, Parkinson's disease or Huntington's
 CC disease), developmental disorders, non-infectious disorders, nervous
 CC system diseases and/or disorders, and autoimmune diseases (e.g.
 CC Addison's disease, haemolytic anaemia, antiphospholipid syndrome, or
 CC allergic encephalomyelitis); the polynucleotides are also useful as
 CC chromosome markers, for chromosome identification, gene therapy, and in
 CC identifying organisms from minute biological samples. This is the amino
 CC acid sequence of a leucine rich repeat containing protein.
 XX
 SQ Sequence 1429 AA;

Query Match 11.9%; Score 759; DB 23; Length 1429;
 Best Local Similarity 37.6%; Pred. No. 2.3e-68;
 Matches 183; Conservative 71; Mismatches 165; Indels 68; Gaps 9;

QY 189 MQPKFAGRLQPPNCRGLCSLAQCLANXK:LFEQQLRKXFLGEGDVSAFLNNMIFQ- 247
 DB 559 LQAPLGRQL-----RDLCSLAAGI:WQKTLFSPDLRKXGLGALITFLKNG:ILQE 612
 QY 248 KOINC-----ERS..... FLATSR 262
 DB 613 HPIPLSYSPHLCFQFFAAMSYLEDEKQPKRKNCTIDLEKTLKAYGNGHGFSASTR 672
 QY 263 FLUGLNEETRSHLEKSLCKWYSHIKVTLCKTCSRAGSDSTLQOCSLEFFSCILYEIC 322
 DB 673 FLUGLSDEGEEMENIFCRISQ--GRNLMGWVPS:Q--LLIQHSLES:HCLETETR 726
 QY 323 REEPLCCALSHFQVIVWSNIAKSWHMVSSFLCKKCRSAQVCHLYGATYSAGDGRARCS 382
 DB 727 NKTETQVMAHFEENGVC VETDMELLVCTCFKFSRHVKMLQLJ-----EGRCHRS 777
 QY 383 AGAHTLLVQLPRTVLLDAYSEHLAAALCTNPN:LELSLYRNALGSRVKLLCOGLRHN 442
 DB 778 TWSPTKWV-LFRWVHTDAYVQVILFSLVKVTRNKLKESLSNLSRSHSAVSLCKTLRRPR 836
 QY 443 CKLQMLR 449
 DB 837 CLETLR 843

RESULT 14
 ABG78472
 IL ABG7969 standard; Protein; 1429 AA.
 XX ABG7969;
 XX
 DT 07 JAN 2003 (first entry)
 DE Human leucine rich repeat domain containing protein #2.
 XX
 KW Leucine rich repeat; nervous system; human; renal disorder; apoptosis;
 KW renal disorder; immune disorder; arthritis; asthma; AIDS;
 KW acquired immunodeficiency syndrome; rheumatoid arthritis;
 KW haematopoietic disorder; metabolic disorder; reproductive disorder;
 KW pulmonary disease; cardiovascular disease; hyperproliferative disorder;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; developmental disorder; autoimmune disease;
 KW Addison's disease; haemolytic anaemia; antiphospholipid syndrome;
 KW allergic encephalomyelitis; gene therapy.
 XX
 CC Homo sapiens.
 XX
 TC M0200274959-A2.
 XX
 PD 26-SEP-2002.
 XX
 FF 20-DEC-2001; 2001W0-US50457.
 XX
 PR 03-JAN-2001; 2001US-259479P.
 XX
 PR 09-JAN-2001; 2001US-260616P.
 XX
 PA (BRIM : BRISTOL-MYERS SQUIBB CO.
 XX
 PI Ramarathan C, Feder J, Mintier G;
 XX
 CR WPI; 2002-750554/91.
 XX
 XX New HLRNSI nucleic acids and polypeptides, useful for preventing,
 PT treating, or ameliorating e.g. renal disorder, immune, haematopoietic,
 PT metabolic, reproductive, pulmonary, cardiovascular or autoimmune
 PT diseases .
 XX
 XX Example 1: Page 392-397; 415pp; English.
 PS
 XX CC The invention describes nucleic acids encoding human leucine-rich repeat
 CC containing proteins expressed in nervous system tissues, HLRNSI. The
 CC HLRNSI polypeptide or the polynucleotide is useful for preventing,

CC treating, or ameliorating a neural disorder or a disorder related to
 CC aberrant apoptosis modulation (either directly or indirectly), renal
 CC disorder, immune disorder (e.g. arthritis, asthma, acquired
 CC immunodeficiency syndrome (AIDS), or rheumatoid arthritis),
 CC haematopoietic, metabolic, reproductive, pulmonary or cardiovascular
 CC diseases, hyperproliferative disorders, neurodegenerative diseases
 CC (e.g. Alzheimer's disease, Parkinson's disease or Huntington's
 CC disease), developmental disorders, non-infectious disorders, nervous
 CC system diseases and/or disorders, and autoimmune diseases (e.g.
 CC Addison's disease, haemolytic anaemia, antiphospholipid syndrome, or
 CC allergic encephalomyelitis); the polynucleotides are also useful as
 CC chromosome markers, for chromosome identification, gene therapy, and in
 CC identifying organisms from minute biological samples. This is the amino
 CC acid sequence of a leucine rich repeat containing protein.
 XX
 SQ Sequence 1429 AA;

Query Match 11.9%; Score 759; DB 23; Length 1429;
 Best Local Similarity 37.6%; Pred. No. 2.3e-68;
 Matches 183; Conservative 71; Mismatches 165; Indels 68; Gaps 9;

QY 379 SLAELIGKSTATPAFIQ:LSRRPPLFILOGVDEPGWVLEQSSSELCLHWSQPQADA 436
 DB 69 LLSNLRKRLPELSLITTRPTALFKLRRLEHPRHVEILGFSEAEKKEYFYKFNAAE 128
 DB 439 LLGSLGKTLPEASFLTANTTALQNLPSLEQARAVEVLGFSESSKKEYFYKFTDEP 498
 QY 129 QASQPNYVRNNEPLFTVCYPLVCMVVCTCLOOLEGCSLRCTSTTTAVVMY:YLSL 188
 DB 499 QAIRAPPVYSNKEKALCLVPMVSWJACTCLNQMKEKEKLTLSKTITLCHYLAQA 558
 QY 189 MQPKFAGRLQPPNCRGLCSLAQCLANXK:LFEQQLRKXGLGEGDVSAFLNNMIFQ- 247
 DB 559 LQAPLGRQL-----RDLCSLAAGI:WQKTLFSPDLRKXGLGALITFLKNG:ILQE 612
 QY 248 KOINC-----ERS..... FLATSR 262
 DB 613 HPIPLSYSPHLCFQFFAAMSYLEDEKQPKRKNCTIDLEKTLKAYGNGHGFSASTR 672
 QY 263 FLUGLNEETRSHLEKSLCKWYSHIKVTLCKTCSRAGSDSTLQOCSLEFFSCILYEIC 322
 DB 673 FLUGLSDEGEEMENIFCRISQ--GRNLMGWVPS:Q--LLIQHSLES:HCLETETR 726
 QY 323 REEPLCCALSHFQVIVWSNIAKSWHMVSSFLCKKCRSAQVCHLYGATYSAGDGRARCS 382
 DB 727 NKTETQVMAHFEENGVC VETDMELLVCTCFKFSRHVKMLQLJ-----EGRCHRS 777
 QY 383 AGAHTLLVQLPRTVLLDAYSEHLAAALCTNPN:LELSLYRNALGSRVKLLCOGLRHN 442
 DB 778 TWSPTKWV-LFRWVHTDAYVQVILFSLVKVTRNKLKESLSNLSRSHSAVSLCKTLRRPR 836
 QY 443 CKLQMLR 449
 DB 837 CLETLR 843

RESULT 14
 ABG78472
 ID ABG78472 standard; Protein; 1429 AA.
 XX ABG78472;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Leucine-rich repeat (LRR) domain containing protein #2.
 XX
 KW Human; human leucine rich repeat; small intestine I; HLRNSI; asthma;
 KW proliferative disorder; gastrointestinal disorder; renal disorder;
 KW neural disorder; reproductive disorder; calcium regulation; apoptosis;
 KW immune system; anaemia; human immune deficiency virus; HIV; cancer;
 KW blood coagulation disorder; autoimmune disorder; allergic reaction;

CC Cardiovascular disorders, and neurological diseases, and for increasing
 CC the organisms' ability to synthesize and/or release pheromones. The
 CC polypeptide may also be used in treating, preventing or ameliorating
 CC pulmonary disease (e.g. chronic obstructive pulmonary disease, allergic
 CC rhinitis, or bronchial hyperresponsiveness), reproductive disease,
 CC haematopoietic disease, platelet disorders (e.g. Bernard-Soulier
 CC syndrome), non-infectious disorders (e.g. innate immunity to bacterial
 CC pathogens, or adaptive immune response), immune and inflammatory
 CC disorders (e.g. systemic lupus erythematosus), cardiovascular diseases
 CC and cancers. HERRM1 nucleic acids may further be used in chromosome
 CC identification or mapping, as a chromosome marker, as molecular weight
 CC markers, as diagnostic probes, in gene therapy, in raising anti-DNA
 CC antibodies, or as antigens for eliciting immune responses.
 XX
 XX Sequence 1429 AA;
 Query Match 31.9%; Score 759; DB 23; Length 1429;
 Best Local Similarity 37.5%; Pred No. 2,20,68;
 Matches 183; Conservative 7; Mismatches 165; Indels 69; Gaps 9;
 QY 9 SMCGLIFSCWPSAPLQSLIRVPERLLFIIDGPRDLKPSFSDPGPWCLQWEEKRPTL 68
 LE 179 SLAELICKGGTATPAPIRQILSRPERLLFIIDGSEPCWULQERSSECLHWSQPACA 438
 QY 69 LNSLIRKKLLPCLSLITRTALEKHLERPRHVELLFSEAEKSEYKYVERNAE 28
 DB 439 LLSGLIGXTILFEAGFLITARTALONLSPSEQARWVEVGFSESSKSEYFYFDER 498
 QY 129 QAGQVENVYDKEPIFTMGFVPLVWVCTCQCLEGGHLLRQTSPTTAVVMYLLSL 188
 DE 499 QAIRAFRLVKSKEWLWCLVFWVSWLACTLWQCKKKEKLTLSKTTTTCLHYLAQA 558
 QY 189 MQPKFGAPLQPPKQKRGCLSAATGLNOKLIFEDQLRKHGLDGEVSAFKNKNIQ- 247
 DB 559 LQAQLGLPOL---RDICSLARGIWOKKTLFSPDLKXHGIGCAISTFLKMGIIQE 612
 QY 248 ---KQING---ERS---KLAISR 262
 DB 613 HPDPLSYSTIHCFQEFPAKSYVLEDEKGRKSHNCTHLEKTLKAYGIHGLFGASTTR 672
 QY 263 FLFGLNEETRSLEKSLCWKVSPIKXPILOWIKSQASDSTQQQSLEFFSCLYEIQ 322
 DE 673 FLGLSLDGEEXENIFICPLSQGRNLMQWVFSIQGLQPSIESLHCLYETR 726
 QY 323 SEEFQCALSHQGVIVVNIANKMRYVAGKQKLPKATVYHLYGATYAGDGERAPCS 382
 DE 727 KATFETQVAHFEKNGC VETMPFLVATFCKPSHRAVACLLI---EGQIHS 777
 QY 383 AGAHTLLVOLPERTVLDAYSEHMAAATNTNLLIELNLYRKALDSKGVKLCQGLRHPN 442
 DE 778 TASPNTVV-LFRWVPTDAYWQILFSVLKATVFLHDELKSNNSLHSAVKSLSLCTIRPR 836
 QY 443 CKLQNR 449
 DE 837 CLETLR 843

Search completed: October 31, 2003, 07:04:04
 Cpu time : 85 secs

GenCore version 3.1.6
Copyright (c) 1993-2003 Computer Ltd.

QM protein - protein search, using sw model

Run on: October 31, 2003, 03:53:49 ; Search time 44 seconds
(without alignment)
338,632 Million cell updates/sec

Title: US-10-028-374-2

Perfect score: 2377

Sequence: 1 MQSATEKSYQDLFSWFE.....HWKMLGGRHPRKQLNR 449

Scoring table: RLOSIM62

Gapop 10.0, Gapext 0.5

Searched: 283309 seqs, 96169662 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum March 1301

Listing first 45 summaries

Database:

- 1: PIR 75.1
- 2: PIR 1.1
- 3: PIR 1.1
- 4: PIR 1.1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB ID	Description
1	759	31.9	1192	hypothetical prote
2	635	22.5	1171	water protein (imp
3	327.5	13.8	483	Maspressin recepr
4	214	9.0	1130	MHC class II trans
5	150.5	6.3	1004	hypothetical prote
6	118.5	4.9	1078	hypothetical prote
7	113	4.8	456	ribonuclease inhib
8	110	4.6	456	ribonuclease inhib
9	109	4.6	388	cyclin B - startis
10	107	4.5	461	ribonuclease-angio
11	106.5	4.5	1232	neuronal apoptosis
12	106	4.5	1020	probable large ATP
13	102	4.3	2110	unknown protein 15
14	99	4.2	506	r complex-associat
15	96.5	4.1	1172	cutin protein - fi
16	95	4.0	2021	serine/threonine k
17	93	4.0	2747	fat facets (fafi) s
18	94.5	4.0	1447	neuronal apoptosis
19	94	4.0	518	glycine hydroxymet
20	93.5	3.9	853	ary hydrocarbon r
21	93	3.9	1277	hypothetical prote
22	91.5	3.8	459	transmembrin FKBP5
23	91.5	3.8	613	hypothetical prote
24	91.5	3.8	613	conserved hypothet
25	91	3.8	1306	hypothetical prote
26	90.5	3.8	797	hypothetical prote
27	90.5	3.8	1402	cell surface prote
28	90	3.8	394	hypothetical prote
29	89.5	3.8	2529	hypothetical prote

30	89	3.7	523	2	A56118	vetispiradiene syn
31	89	3.7	534	2	T05544	hypothetical prote
32	89	3.7	1204	2	D71490	probable exodeoxyr
33	89	3.7	4861	2	S71752	giant protein p619
34	88.5	3.7	564	2	E75394	ABC transporter th
35	88.5	3.7	921	2	D95113	malt regulatory pr
36	88	3.7	389	2	A42268	heterocyst envelo
37	88	3.7	585	2	S48951	hypothetical prote
38	88	3.7	611	2	C89243	protein F28Q1.3 (1
39	88	3.7	631	2	T21471	hypothetical prote
40	88	3.7	1057	2	T46193	hypothetical prote
41	87.5	3.7	453	2	S74503	hypothetical prote
42	87.5	3.7	1023	2	A71643	ATP-dependent dscN
43	87	3.7	312	2	B97746	hypothetical prote
44	87	3.7	512	2	T39422	probable chromatin
45	87	3.7	843	2	T01418	hypothetical prote

ALIGNMENTS

RESULT 1

T17255

hypothetical protein DKFZps86G1822.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text_change 15 Oct-1999

C/Accession: T17255

R:Koehrer, K., Beyer, A., Vewes, H.W., Gassenhuber, G., Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z19122

A:Accession: T17255

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1192 (KOE)

A:Cross-references: EMBL:AL117470

A:Experimental source: adult uterus; clone DKFZps86G1822

C:Genetics:

A:Note: DKFZps86G1822.1

Query Match: 11.3%; Score 759; DB 2; Length 1192;
Best Local Similarity 37.6%; Pred. No. 6.3e-57;
Matches 183; Conservative 71; Mismatches 165; Indels 65; Gaps 9;

Cy	9	SMGLFSCWFRSAPLQENRYPERLLYIDFELKPSFDQSPQPCWLOWEKRTPEL 68
Db	98	SLAEILHTFHTATPATIQTLSRPERLLFIDGVDPGQVWLPQSPSELCKRSQCPADA 157
Cy	69	LKASLIERKILPRLSLLITRTIACEKILPILPDRHVEILGFSEAEKKEYFYKHNAE 128
Db	158	LASLIGKNTIDPASFTTARTALQNLPSLRQARWVEVLFSSSRKEYFYRTDTR 217
Cy	129	QAGQVFNVYVNRPELFTYGFVPLVGVVCTCQQDLEGGLRQTSRTTAYVMYLKSL 198
Db	218	CAIRAFELKSNKELNALCVPMWVSLACTCMQMKREKLTLSKTTTLCILHYLQA 277
Cy	189	MQQKPGAPRIQPPFNKSGCSLAACQWQKILFEDQDKRKHGLDGVSAFLNNYIQ- 247
Db	278	LAQCP-SPQL --- PCLSLAAGTGWKTKTSPFDLKRKHGLGGA-ISTFLKYGILQE 331
Cy	248	-----KQINC-----ERS-----EALITSR 262
Db	332	HPILKSYSFILHCFQFFFAKSYVLEDEKGRKUNCIIDLEKLEAYGIRHGLFASITR 391
Cy	263	FILGLINEETFSLEKSLCWKYSPIHKDLQWISKAOSDGTIQQSLEFFSCYEQ 322
Db	392	FELGLLDFEGEREKEMIFNCRISQ--GRKQWQVPSLQ---LLQPHSLSLHCLCYETR 445
Cy	323	EEEFICALSHFPVIVVNSIAKKVHNSFLCKRCSQVHLVGYATYSADCEPRACS 382
Db	446	KKFLQWMAFFENKVC VETVELVATFTFRFSRVKKLQLI-----EQRCHRS 496
Cy	193	AGAPTLLVQVHPTVLDVAYSHLAAGATNPKLVELSYRNALDSRGVKLCQGLRHPK 442

DQ 497 TWSPYVW-LPRKVVTDAYWQILPSVI KVTNRLRELISQSLSHSAVKSCKTKTLRPR 555
 QY 443 CKQONLR 449
 DB 556 CUEETLR 562
 RESULT 2
 A59000
 mter protein [imported] - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 01-Aug-2001 #sequence, revision: 01 Aug 2001 #text change 01-Aug-2001
 C:Accession: A59000
 K:Long, Z.B.; Nelson, J.M.
 Endocrinology 140, 3726-3726, 1999
 A:Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature ov
 A:Reference number: A59000; MIMD:99360614; PMID:10433232
 A:Accession: A59000
 A:Status: preliminary
 A:Molecule type: cDNA
 A:Residues: 1-111 <KUR>
 A:Cross-references: GB:AF074016; NID:9402692; PID:AA061142.1; PID:95802698
 C:Genetics:
 A:Gene: Muter
 Query Match 22.5%; Score 51.7; DB 2; Length 111;
 Best Local Similarity 29.1%; Pred. No. 1.2e-37;
 Matches 143; Conservative 85; Mismatches 20; Indels 56; Gaps 6;
 QY 9 SYGCLIFSCWPEPSAPQELIRVPEPLRIINPFLAPLAFDPQGPWLNWEEKRPTEL 68
 DB 241 SLACIARCPDSWNVTKMS-PEPLAVIADIKMSVLCHEHMTSRWNKEPPI: 308
 QY 69 LANSIRKALPELSLITETETALEKRLHRLHPRKELIAPSRAPKEVEYKYPHAE 126
 DB 301 LMSILKALIPQSLITRATGLEKLSMVSRLYLIVELASASRSQVLENISNE 360
 QY 129 CACQENYVRDEPLTMEVEVRLVAVTCTCCQCEPEPLIPUTSRITVAV...MYL 185
 DB 361 DRTOVSHSGIERHQLPQCGANVSILVGEALGLKRIQAKHLRQGLTGLYATLVPHQ 420
 QY 196 LSNQPKGAPRLCPDPPNGKCSLAANDKWKIIFRQELERHQLSGEDVSAFLNNKI 245
 DB 421 LTLKRSQALSQESQITLVGLVYALSNVWNRVFNGLNLSKSESEILAFEMNZ 480
 QY 246 PQDNDERS...FIAL... 259
 DB 481 LLOVHNSQKYVFSLSLQCFENALVYVLELLEWNSJKEQFENQVMEVARTYTRL 540
 QY 260 TSREPLGLAREETSHLEKALWVNSDINVYLLQWILQKACQUGSTLQSGSLEFFSC 317
 DB 541 LKXKRLFGSKNKILKTLTEVFEVTVITVTKKCHVSLACQVNGTSPKCTLEAFYC 600
 QY 318 LYETQREITQALSFPQVTVENIASRNEHVSPFLKSRKQAVIHLVGYATSGADGD 377
 DB 601 LRFSCQSEFVGKALRFQEWLLINQAVELKVSAYLQKQKRAIVLIRQLSVQNT 659
 QY 378 PARCSAGMTLLVQDERTVLLANSEHAAALCTNPKIIELSLERNMAGRGVLLQGS 437
 DB 660 LELCP...VTVYQEQCPLEWKNQFYSVGLADPKKELELQGSLSQRAWKILCLE 715
 QY 438 LRHPNCKLQNL 448
 DB 716 LRNQSRIOQL 726
 RESULT 3
 S2780
 Nasopressin receptor rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Apr-1993 #sequence, revision: 10-Apr-1993 #text change 05-Nov-1999
 C:Accession: S2780
 R:Herrera, V.L.; Ruiz-Opazo, N.

submitted to the EMBL Data Library, February 1992
 A:Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.
 A:Reference number: S2780
 A:Accession: S2780
 A:Molecule type: mRNA
 A:Residues: 1-483 <HER>
 A:Cross-references: ENEL:82518; NID:920285; PID:AA036231.1; PID:9202856
 C:Genetics:
 A:Gene: AVP

Query Match 11.8%; Score 32.75; DB 2; Length 483;
 Best Local Similarity 27.9%; Pred. No. 1.2e-20;
 Matches 109; Conservative 53; Mismatches 123; Indels 155; Gaps 11;
 QY 164 LGGGLRQTSPTTAVVLYLLSYLSPKPGARLPQPPNQRGLCSLAADGLMKNKILF 222
 DB 1 MELGRLSRTSKTTSVYLFF-TGMLKSAGTNSPRVQ--GELRMLCLRLAREGILKHQAF 58
 QY 223 EEOELR-KHGLQEDVSA-----PLNNK----- 245
 DB 59 SEXTLEALKLKGSGVQVTFSLRKELPQVLETVVTVYFIDQSFOEPLAALSYLLDAEGA 118
 QY 246 -----PCKDINTE--RSFLATSRPFGLNSETSRSHLEKSLCWKVSPIKVELLQW 295
 DB 119 PENSASISQVMLNSDAHLRCHLALTFPFGILSTERIDIONHFGCVVPGVKCTLEK 178
 QY 296 IQSKAQSDST----- 330
 DB 179 VQGSQKQVATVGAERKDELKDFEAEEREEEREEELNFGLELLYCLYETQEDDFVQA 218
 QY 331 LSHFQVTVSN: ASKVEHNVSSFCARCESQVHLVGYATYSADGEDRA:----- 380
 DB 239 LSSAFEMVLRVRLTMDLEVLSCVCCDQCALRVSCGLVAAKEKKKKKSMNRJL 299
 QY 481 CSAGAR:TLVQGLPRTV-----LDDAYSEHLAAALTNPKLIELS 420
 DB 299 GSSTCKQPHASLRPLSEAMITQCGSLITLSHKIPDAVCRLSELKALVAPSLREL 358
 QY 421 LYNALSRVRLICQGLRHPNCKLQNR 440
 DB 359 LQGNRTEAGRLISQSLAMPKCKVQTLR 387

RESULT 4
 A48843
 MHC class II transactivator human
 N:Alternate names: AIIA
 C:Species: Homo sapiens (man)
 C:Date: 06-Jan-1995 #sequence, revision: 06 Jan-1995 #text change 08-Oct-1999
 C:Accession: A48843
 R:Steinme, V.; Otten, J.A.; Zufferey, M.; Mach, B.
 Cell 75, 135-146, 1993
 A:Title: Complementat ion cloning of an MHC class II transactivator mutated in heredit
 A:Reference number: A48843; MIMD:94026536; PMID:8402893
 A:Accession: A48843
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-110 <SLE>
 A:Cross-references: GRX474101; NID:9414112; PID:CAAS2354.1; PID:9414113

Query Match 9.0%; Score 21.4; DB 2; Length 1130;
 Best Local Similarity 24.3%; Pred. No. 6.6e-10;
 Matches 118; Conservative 61; Mismatches 187; Indels 120; Gaps 23;
 QY 1 MNQATGCSMQLFSCWPEPSAPQES---LIRVPERLJFIIDGFDELKPS---PHUPQ 51
 DB 456 LKRPDQATGILQLFSGQPLVAADVEFSLKRFDRVLLILDADFEELEAODGFLHSC 515
 QY 54 GPWTLCKEKEEETEL--LINSJIRKLLPELSLITTRP-----TALKKLELLEHPRV 106
 DB 516 GP--APAEFSKMLLAGFQKALRGCTTLTARTPRGLVQSLSKADALF----- 564
 QY 107 ETLQREAEERKEVEYKYPHNA---EQAGGVFNVRDNEPLFTMCEVPLVYCVYCTCQCC 163

```

Db 565 ELSDPSMECAQAYVYRYESSOMTEHQDPAALTLKRMFLLSHSHSPCLCRVCCSEAL 624
QY 164 LECGGLLQTSCTTAVVYLLVLYLQYQKAPAPLQPPHQKQGLCSLAADGLKCKKL 221
Db 645 ELGECALPLSFLTSHVGLD-----GPAALCSPPALASLAKLAEELGRHGSTC 675
QY 224 FEQD-----DURKIGS-----QDEVSAP-----LKNYFQKQINC 252
Db 616 QEDQPSADVRTWAKGLVYUHFRAESLAFSELLQFELGALKALASLSEIKDEL-- 733
QY 253 ENSFLALT-----SAPFLHLSNEFSSHFKKGLQKVKVSHKMLLOWIQS 298
Db 734 PQYALTPRKERYDNWLEQVPEFADLIFQFARVLSALLGPSAAASVDRKQVLA 790
QY 299 --KAQSDGST:QGSLSDEFFSLVLDCEEFQCALSHHFQVYVNSNTASKMEHWSSFC 355
Db 791 RYKRLQQTQPARQLLELHCAHAEAGTQCH-----VQDELPGD-----SFLG 837
QY 356 KRCSAQVHLHYGATYSAGDE-----PARCSAGAHITL-----VQLPERTVLL 399
Db 839 TRLTPTDA-HVGLKALEMAGGCSLQJSTGLCTGLSLVGLSCTVPRFRAALSCTVALW 896
QY 400 DAYSEH 405
Db 897 ESILPQH 502

RESULT 5
T31665
Hypothetical protein GUS1.4 - sea squirt (Cliona intestinalis)
C:Species: Clona intestinalis
C:Date: 29 Oct 1999 #sequence revision 29 Oct 1999 #text_change 20 Jun 2000
C:Accession: T31665
R:Bird, A.P.; Clark, V.J.; Jones, S.J.; Lettgeb, S.; Tennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z1050
A:Accession: T31665
A:Status: preliminary; translated from GB/EMBL/DBCB
A:Molecule type: DNA
A:Residues: 1:1075 <R>
A:Cross-references: EMBL:Z282904; P:DN:CA802544.1
C:Genetics:
A:Introns: 30/1; 539/1; 593/1; 662/3; 724/1; 769/1; 817/2; 841/2; 960/1

Query Match 6.43; Score 167.5; DB 2; Length 1075;
Best Local Similarity 19.83; Pred. No. 5.16;
Matches 99; Conservative 79; Mismatches 126; Indels 145; Gaps 23;

QY 20 EESAPQJELIRVPERGLTFLIDFHIEKIFPHILQYQWTLKHEKRTIELLSLRKK 77
Db 132 ELSKALITFLNNPNVAFFDGLIFASTNEFARPHICKLDGKSPVDIMKN-LFNDA 188
QY 78 ILPELSLITTPPTALEKHLRLEHFRHVEILGSEAPREYKYFPHNAEQAGQVNVY 137
Db 189 LPAKIVVTSTPDMFNLOHCYPTPSFEVLGSLAEAKNL-----GTCLGCKYPA 242
QY 138 R--DNEP-LFTMCPVPLVQVYVTCGQCGLEGGILLRQTSRTTAVVYLLSLMQ--- 190
Db 243 KKLIDQPNLAHLQYLPNFILVFCLLN-EGSDI-----KMTQVLFSTNFRVESH 296
QY 191 -----PKGAPRIOPPNCRGSLAAGLNKXKILFEQQLKHKCLDGE----- 235
Db 297 LKGEVPLDKVGAEMVK-----LACLAYKHGCKKLVFEKTFDQVVKLADEVYTNFT 349
QY 236 --DVSALNNM:FQDKINCERSFL-----LQVLEKLVKLVKLVKLVKLVKLVK 458
Db 350 YVDLSSGIRIKILEN--KSYFHHIQQEYAAVYLVLYLVSYREFEQKTIPTDQWS 406
QY 259 JTSRLEFGUN-----KSTPSHLEKSLQKVSUHPHCKMDLQ W: 246
Db 407 VVVKFMGICNPAYKQLIKLIPFATNIDYEEKKFEESYWEISAKSECCLPREFSWL 466

```

```

QY 297 QSKAQSDGST:QGSLSDEFFSCL-----YEIQEEEFIOCALSHF----- 334
Db 467 HE--INCDET--SKRFKQCLFVGLKAVGVHKLPEVKOLVYALKSFTKPKHLRNSWT 520
QY 335 --QV-----VYKNTASKMEHWSSFCUKRCSAQVHLHYGATYSAD 374
Db 521 TTTEVLETGRIHETITTTTFVYNNKEMQSDLM-----ELLQHL-----D 563
QY 375 GEDRARCAGAHITLVCLEPRTVALAVSEHAAALCINPNILIELSY 422
Db 564 AVEKLEF-----QVTHLSYNSLSINORSNKIOLDJW 599

RESULT 6
T31668
Hypothetical protein GUS1.5 - sea squirt (Clona intestinalis)
C:Species: Clona intestinalis
C:Date: 29 Oct 1999 #sequence revision 29 Oct 1999 #text_change 20 Jun 2000
C:Accession: T31668
R:Bird, A.P.; Clark, V.J.; Jones, S.J.; Lettgeb, S.; Tennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z21050
A:Accession: T31668
A:Status: preliminary; translated from GB/EMBL/DBCB
A:Molecule type: DNA
A:Residues: 1:1075 <R>
A:Cross-references: EMBL:Z282904; P:DN:CA802544.1
C:Genetics:
A:Introns: 48/1; 557/1; 611/3; 685/2; 727/2; 739/1; 779/3; 820/3; 853/2; 914/2; 993/1

Query Match 4.91; Score 116.5; DB 2; Length 1075;
Best Local Similarity 19.53; Pred. No. 5.16;
Matches 98; Conservative 76; Mismatches 203; Indels 125; Gaps 19;

QY 20 EESAPQJELIRVPERGLTFLIDFHIEKIFPHILQYQWTLKHEKRTIELLSLRKK 77
Db 132 ELSKALITFLNNPNVAFFDGLIFASTNEFARPHICKLDGKSPVDIMKN-LFNDA 208
QY 40 PELSLSLITTPPTALEKHLRLEHFRHVEILGSEAPREYKYFPHNAEQAGQVNVYR 139
Db 209 PVAKIVVTSTHMYKILHPDYRPTPSFEVLGSLAEAKNL-----GTCLGCKYPA 262
QY 139 -DNEP-LFTMCPVPLVQVYVTCGQCGLEGGILLRQTSRTTAVVYLLSLMQ--- 190
Db 263 ILQQPNLAHLQYLPNFILVFCLLN-EGSDI-----KMTQVLFSTNFRVESH 316
QY 191 --PKGAPRIOPPNCRGSLAAGLNKXKILFEQQLKHKCLDGE----- 226
Db 317 GAVP-LKVGAEVVK-----LAPLAYKGLQKLVFEKTFDQVVKLADEVYTNFT 369
QY 227 -----LRKFLDGEHVSAPLVNVIPOKINCERSFLALTSR----- 262
Db 370 QKSSGIRIKILEN--KSYFHHIQQEYAAVYLVLYLVSYREFEQKTIPTDQWS 428
QY 263 PLSGLINEETSHLEKSLQKVSUHPHCKMDLQ-----HMKDGLLOWIQSAGDST 315
Db 429 PFGICNPAPKQCK-----LVFFATN:KYEEKELM--VPMMSLSMARGDLIRF 461
QY 316 SCLYATQEEETFOCALSHFQV-----IVWSNIASKYEHMVSSFCUKRCSAQV 370
Db 492 GWLHVNLGERSKPFEDY:PVGLKVDAPKHLSEVKDLV--YALKSFTKPKHLRSDSV 539
QY 371 YSADIEDPARSAGAH-----LLVQLPE-----RTVLLDAYSE 404
Db 540 TTEVLETIRGVHGTITTTTFVYNNIEMKQSLMELLHLHLDAMEELRFRVTKLSY 599
QY 405 KLAALCINPNILIELSYNAL 426
Db 600 RLSNALNCSNKIQLVILKRL 621

RESULT 7
A31857

```


[illegible]

Search completed: October 31, 2003, 07:02:37
Job time : 5: secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	2334	99.2	1062	1	PV47 HUMAN	P5904F homo sapien
2	1157.5	48.7	1034	1	C511 HUMAN	Q94620 homo sapien
3	1146	48.2	1033	1	C515 MOUSE	Q94458 mus muscula
4	725	41.9	1473	1	NALI HUMAN	Q50002 homo sapien
5	674.5	28.4	892	1	PV45 HUMAN	P59041 homo sapien
6	671.5	28.2	843	1	PV46 MOUSE	Q14852 mus muscula
7	655	27.6	854	1	PV48 RAT	Q61035 rattus norv
8	652.5	27.5	994	1	NALI HUMAN	Q60622 homo sapien
9	564	23.7	1562	1	NALI2 HUMAN	Q94522 homo sapien
10	563.5	23.7	980	1	PV43 HUMAN	Q94594 homo sapien
11	556.5	23.4	1200	1	NALI2 HUMAN	P59047 homo sapien
12	535	22.5	1111	1	NALI5 MOUSE	Q94175 mus muscula
13	428.5	16.0	1033	1	PV46 HUMAN	P59045 homo sapien
14	261.5	11.0	1040	1	CARF HUMAN	Q94529 homo sapien
15	261	11.0	953	1	CAR4 HUMAN	Q94239 homo sapien
16	253	10.6	1020	1	CARF MOUSE	Q94120 mus muscula
17	233	9.8	953	1	CAR4 MOUSE	Q94120 mus muscula
18	214	9.3	1130	1	C27A HUMAN	Q94120 mus muscula
19	174	7.3	1155	1	C27A MOUSE	P33176 homo sapien
20	117.5	4.9	1403	1	B1B1 MOUSE	P79621 mus muscula
21	113	4.8	1456	1	B1B1 PID	Q94045 mus muscula
22	110	4.6	456	1	RINI RAT	P13275 sus scrofa
23	109	4.6	388	1	CG4F MARGE	P23115 rattus norv
24	107	4.5	460	1	RINI HUMAN	P13269 marthasteri
25	106.5	4.5	1403	1	B1B1 HUMAN	Q13075 homo sapien
26	105.5	4.4	1824	1	CAR6 HUMAN	Q94474 homo sapien
27	105.5	4.4	1403	1	B1B1 MOUSE	Q94016 mus muscula
28	102.5	4.2	1402	1	B1B1 MOUSE	Q94203 mus muscula
29	98.5	4.1	1402	1	B1B1 MOUSE	Q94103 mus muscula
30	98.5	4.1	1403	1	B1B1 MOUSE	Q94106 mus muscula
31	96.5	4.1	1172	1	SM2 SCRPO	P41001 schizosacch
32	95	4.0	2178	1	FAP PROME	P54824 drosophila
33	94.5	4.0	1447	1	B1B1 MOUSE	Q94044 mus muscula

CC Event/Alternative splicing, Named isoforms: 3;
 CC Name=1; Synonyms=1;
 CC IsoId=P59046-1; Sequence=Displayed;
 CC Name=2; Synonyms=1;
 CC IsoId=P59046-2; Sequence=VSP_005824;
 CC Name=3; Synonyms=1;
 CC IsoId=P59046-3; Sequence=VSP_005823;
 CC TISSUE SPECIFICITY: Detected only in peripheral blood leukocytes,
 CC predominantly in eosinophils and granulocytes, and at lower levels
 CC in monocytes;
 CC SIMILARITY: Contains 1 CAPK domain;
 CC SIMILARITY: Contains 8 leucine-rich (LRR) repeats;
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the ExPb Institute on
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/submit/submit.html>)
 CC or send an email to license@ebi.ac.uk;
 CC
 DR EMBL: AY095146; AAM13227.1;
 DR EMBL: AY116203; AAM75142.1;
 DR EMBL: AY116203; AAM75143.1;
 DR EMBL: AY116203; AAM75144.1;
 DR EMBL: BC028069; AAM78069.1;
 DR InterPro: IPR001611; LRR;
 DR InterPro: IPR007091; LRR_RNCH;
 DR InterPro: IPR003590; LRR_RNCH_sub;
 DR InterPro: IPR007111; NACHT_NFase;
 DR InterPro: IPR004320; NACHT_DAPIN_dom;
 DR Pfam: PF00560; LRR; 2;
 DR Pfam: PF02758; NACHT_DAPIN; 1;
 DR PRINTS: PR000117; LEUC1CHRPT;
 DR PROSITE: PS00368; LRR; 1;
 DR PROSITE: PS00824; DAPIN; 1;
 DR PROSITE: PS00837; NACHT; 1;
 KW ATP-binding; leucine-rich repeat; Repeat; Alternative splicing;
 FT DOMAIN 1 95 DAPIN;
 FT REPEAT 211 528 NACHT;
 FT REPEAT 713 736 LRR 1;
 FT REPEAT 742 765 LRR 2;
 FT REPEAT 770 792 LRR 3;
 FT REPEAT 799 822 LRR 4;
 FT REPEAT 827 850 LRR 5;
 FT REPEAT 894 907 LRR 6;
 FT REPEAT 941 969 LRR 7;
 FT REPEAT 978 1021 LRR 8;
 FT N-TERM 217 224 ATP IDENTICAL;
 FT VARSPLIC 963 974 Missing (in RefSeq)
 FT VARSPLIC 977 1012 Missing (in RefSeq)
 FT CONFLICT 632 692 Missing (in RefSeq)
 SQ SEQUENCE 1062 AA; 120328 MW; 0A8B1C8F1B49F CRO54;
 Query Match 98.2%; Score 2334; Pos 1; Length 1062;
 Best Local Similarity 90.7%; Pos 1; Length 168;
 Matches 449; Conservative 0; Mismatches 0; Indels 46; Gaps 2;
 QY 1 MNQSTECNQLIPSCWPEPSAPLOELIRVPERLLEIDGFDLKPSPHSPQGWCLCW 60
 RB 1 MNQSTECNQLIPSCWPEPSAPLOELIRVPERLLEIDGFDLKPSPHSPQGWCLCW 314
 QY 61 EKKRTEELNSLKKLPESLLELTPALEKHLHLEHFRVRLGSESEKREYF 120
 DB 15 EKKRTEELNSLKKLPESLLELTPALEKHLHLEHFRVRLGSESEKREYF 374
 QY 121 KYKFNACAGOVNVRNDEPFCVGVVGVWVWVTCQQLGGLPQTSRTTAV 180
 DB 175 KYKFNACAGOVNVRNDEPFCVGVVGVWVWVTCQQLGGLPQTSRTTAV 434

QY 121 YMLYLLSLKQPKSARLQPPNPGKLSAADGLWKNQKILFEEQDLKXKGLDGDVSF 240
 DB 435 YMLYLLSLKQPKSARLQPPNPGKLSAADGLWKNQKILFEEQDLKXKGLDGDVSF 494
 QY 241 LKXNIFQKIDNG... ..ERS 255
 DB 495 LKXNIFQKIDNG... ..ERS 554
 QY 256 FALTSRFLFG... ..ERS 315
 DB 555 FALTSRFLFG... ..ERS 614
 QY 316 SCVLECEEEF... ..ERS 375
 DB 615 SCVLECEEEF... ..ERS 674
 QY 376 EGRAPCSAHTLV... ..ERS 434
 DB 675 EGRAPCSAHTLV... ..ERS 734
 QY 435 CCGRHPNCKQNL 449
 DB 735 CCGRHPNCKQNL 743
 RESULT 2
 CIS: HUMAN STANDARD: PRT: 1514 AA.
 AC 096P23; C75434; ONTOW; CATEU9; Q6XK99;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cold autoimmune syndrome 1 protein (Cryoaprin) (NACHT-, LRR- and
 DE PVD-containing protein 3) (PVDK-containing APAF1-like protein 1)
 DE Angiotensin/vasopressin receptor AL/AVP-like;
 DE CAS1 OR NALP3 OR PYPAF1;
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo;
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOPFORMS 1 AND 2), VARIANTS FCAS MET-196; VAL 419
 RP AND GAY-627, AND VARIANT MWS VAL-352;
 RX MEDLINE:21547523; PubMed:11657797;
 RA Hoffman R.M., Mueller J.L., Brink E.H., Waudet A.A.,
 RA Kolodner R.D.;
 RT "Mutation of a new gene encoding a putative pyrin-like protein causes
 RT familial cold autoinflammatory syndrome and Muckle-Wells syndrome";
 RL Nat. Genet. 29:101-105 (2001);
 RN [2]
 RP SEQUENCE FROM N.A. (ISOPFORM 2);
 RX MEDLINE:2193506; PubMed:11786556;
 RA Manji G.A., Wang L., Geddes B.J., Brown M., Merriam S., Al-Garawi A.,
 RA Mak S., Gora J.M., Briskin M., Curtan M., Cao J., DiStefano P.S.,
 RA Bertin J.;
 RT "PYPAF1: a PYPIN-containing APAF1 like protein that assembles with ASC
 RT and activates NF-kB";
 RL J. Biol. Chem. 277:11570-11575 (2002);
 RN [3]
 RP SEQUENCE FROM N.A. (ISOPFORMS 1, 2 AND 3), VARIANT MWS MET-198, AND
 RP VARIANTS FCAS/MWS TRP-260 AND PRO-305;
 RX MEDLINE:22242314; PubMed:12355453;
 RA Aganica E., Martinot F., Hawkins P.N., Ross C.B., Swar D.C.,
 RA Booth B.R., Lachmann H.J., Gaudet R., Woc P., Feigbery C.,
 RA Cotter F.E., Thorne N., Hiltan G.A., Techoop J., McDermott M.F.;
 RT "Association of mutations in the NALP3/CAS1/PYPAF1 gene with a broad
 RT phenotype including recurrent fever, cold sensitivity, sensorineu-
 RT deafness, and AA amyloidosis";
 RL Arthritis Rheum. 46:2445-2452 (2002);
 RN [4]
 RP SEQUENCE OF 301-1514 FROM N.A. (ISOPFORM 1);
 RX TISSUE=BLIND;
 RX MEDLINE:23493377; PubMed:11042152;

QY 159 NQKFGA--PRLQPPNPGGLCSLAAGLWCKKTLFREQDLPKHGJGDGV-SAFIKV-- 243
 DB 427 ISSAPVAGPRLQ--GDLRLCLRLAREGVLRRAQPAEKELEQELKXGKVTUFSKKE 484
 QY 244 --NIFQKDNCE----- 261
 DB 485 LQGVLETETVTCFIDQSCFELAAI SYLEEDGGVPRFMAGVGTLLRQDQPHSHLVLT 544
 QY 262 RFLGGLNEERSHLEKSLCWVSHVHKKLLQWTCSEAQ----- 302
 DB 545 RFLGGLSARMRDTERHFGGVSRVQCEALRWQGGQCGCPVAREVTEGAGLEDE 604
 QY 303 DOSTLOQSS-----JEFFSCVETQSEEFQQAASHFOVAVVSNIA-SKXEHVSSFCCLK 356
 DB 605 RPEEEEGEERNPYLELYCYLQEDAFVQALCFPELALQVRGCRDVAVLSVCP 664
 QY 357 RCRSAQVLEJ-----YCATYSADTEEARCSAGATL----- 388
 DB 665 CPGACALELSCRLVAACERKKKSLGRKLCASLGGSSSGTQKLPASJLHPLFOAMT 724
 QY 389 --LQVLPRTV-----LDPAYSEHIAAACTTPNPLHLSLYRNALNSPOVKLLCCGRHPN 442
 DB 725 DPKCHLSLTLSHCKLPDAVCRDLSEALRAAPATTELCGLMNLSPFAGLAVLSGLAWPO 784
 QY 443 CKLQNLRL 449
 DB 785 CRVQTVR 791

RESULT 6
 PYAF5 MOUSE
 ID PYAF5 MOUSE STANDARD; RT, 843 AA.
 AC Q9:AS2: Q9RGA4;
 DT 28-FEB-2003 (rel. 41, Created)
 DT 28-FEB-2003 (rel. 41, Last sequence update)
 DT 15-SEP-2003 (rel. 42, Last annotation update)
 CR PYAF5 containing APAF5-like protein & like
 CS PYAF5.
 CS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxId=10092;
 RN [1]
 RF CONCEPTUAL TRANSLATION OF 1-287.
 RA Rinz J.;
 RC Unpublished observations (FEB 2003).
 RN [2]
 RP SEQUENCE OF 288 843 FROM N.A.
 RC TISSUE=Kidney;
 EX MEDLINE=22368257; PubMed=12477312;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Berge J.D.,
 RA Klausner R.D., Collins F.S., Wagner G., Sherman C.W., Schuler G.C.,
 RA Alfesul S.F., Zeeberg B., Buotow K.H., Schaefer C.F., Abat N.A.,
 RA Hopkins R.F., Jordan H., Moore T., Max A.M., Wang G., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin S.W., Hong J.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T., Scheetz T.E.,
 RA Brownstein M.C., Usdin T.B., Toshiyuki S., Cattini P., Prance C.,
 RA Raba S.S., Loquellaro N.A., Peters G.G., Altman R.D., Mullany S.J.,
 RA Rosak S.A., McEwan P.C., McKernan K.J., Vitek J.A., Gamarathne P.H.,
 RA Richards S., Worley K.C., Hale S., Guella A.M., Gay L.J., Halyk S.W.,
 RA Villalon D.K., Wuzny D.M., Sodersten E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettner M., Yadan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Bakesley R.W., Touchman C.W., Green E.D., Jackson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M., Skalska U., Skallus D.E.,
 RA Scherch A., Schein J.E., Jones S.C.M., Vavra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences".
 RJ Proc. Natl. Acad. Sci. U.S.A. 99:16693-16693 (2002).
 RN [3]
 RP IDENTIFICATION OF MAXVALLIAN ORTHOLOGS OF PYAF5.
 RA Albrecht M., Domingues F.S., Schreiber S., Bengauer T.,

*Identification of mammalian orthologs associates PYAF5 with distinct
 functional roles.*
 FEBS Lett. 538:173-177 (2003).
 CC 1- FUNCTION: May mediate activation of CASP1 via ASC and promote
 activation of NF-kappaB (By similarity).
 CC 1- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
 CC 1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC 1- SIMILARITY: Contains 1 DAPIN domain.
 CC 1- SIMILARITY: Contains 1 NACHT domain.
 CC 1- SIMILARITY: Contains 1 leucine-rich (LRR) repeats
 CC 1- CAUTION: the N-terminus was extended using ESTs and genomic
 sequences, in analogy to ortholog sequences.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: BC013519; AAF13519.1; ALT_INIT.
 CC EMBL: BC031139; AAH31139.1; ALT_INIT.
 CC MG: MG:214:99C; Pydafs.
 CC InterPro: IPR001511; LRR.
 CC InterPro: IPR007391; LRR_RNinh.
 CC Pfam: PF05560; LRR; 1.
 CC PROSITE: PS00824; DAPIN; 1.
 CC PROSITE: PS00837; NACHT; 1.
 CC ATP-binding; Leucine-rich repeat; Repeat.
 CC DOMAIN; 1; 102 DAPIN.
 CC DOMAIN; 164 485 NACHT.
 CC FT DOMAIN; 569 585 ASP/GLU-RICH.
 CC FT DOMAIN; 654 66 POLY-LYS.
 CC FT REPEAT; 434 459 LRR 1.
 CC FT REPEAT; 609 632 LRR 2.
 CC FT REPEAT; 811 834 LRR 3.
 CC FT NP_BIND; 174 181 ATP (POTENTIAL).
 CC SQ SEQUENCE 843 AA; 94592 MW; 35FE7A766A47D951 CRC64;
 Query Match 28.2%; Score 671.5; DB 1; Length 843;
 Best Local Similarity 32.5%; Pred. No. 1; 6e-48;
 Matches 178; Conservative 86; Mismatches 168; Indels 115; Gaps 15.
 QY 9 SNGDLIFGWPFEAPAGELIKVHPLAFITGDFELKPSHPDQGPWCLQWKEKRPTEL 68
 DB 220 SLADVLLGTPRAVWPKPLAQHRLFLTGADQLPTLPSSSEATPC ---KPLEA 273
 QY 69 ---GNSLIRKKLLPFLSLTTRPTAI EKLHRLLEHPRHVEILGPSEAKRKYFYKY 123
 DB 274 TSGRVLSLSSGCELRFGARJLVTRHAATGQGRCLSPQCAEIRGFSDDKKKYFFRF 333
 QY 124 FINKAQAQVNVYVRCNEPLEFTMQFVTLVGVVYVTCLOQLQEGGLLRQTSRTTAVYML 183
 DB 334 FDEPKAEPAVAVPKENETLPAICVPFVCMVCTVLOQLLEGLRLSRTSKTTSVYLA 393
 QY 184 YLLSLNQK-PGAPRLQPPNPGGLCSLAAGLWCKKTLFREQDLPKHGJGDGV SAF 243
 DB 394 FTSKXSNAGTNGPRVQ--GEJTLCLRLAREGLDHHKAESEEDLEKLRGSGVQTF 451
 QY 241 LNMWIF-----CKDNCE--RSP 256
 DB 452 LMKKEIPGVLTETVTCFIDQSCFELAAI SYLEAERTPTGAGVOKLNSDAELRGH 511
 QY 257 LAUTGRFGLNLNERTSHLEKSLCWVSHVHKKLLQWTCSEAQ----- 304
 DB 512 LAUTGRFGLNLNERTSHLEKSLCWVSHVHKKLLQWTCSEAQ----- 571
 QY 305 ---STIQSSLEFFSCVETQSEEFQQAASHFOVAVVSNIA-SKXEHVSS 351
 DB 572 QLEDAHLEEBEEDENFG-LESLCYLQEDAFVQALCFPELALQVRGCRDVAVLSVCP 630
 QY 354 SCLXPRQSAQVILRYGATYSADTEEARCSAGATL-----GATLLVQLP----- 393

631 NYVCCDDGQALRVSCGLVAAREKKKKSLVKEKSGSOSTKKQPPVSMRELCEMT 690
394 --- ERTVLG-----PAYSEHAAALATPNDLIESLYNALSGROVKLCCGGLRHN 442
691 TPQCHLSVILSHQRIPDAVCRDGLFAKVAFALESLVSLTETGFLCEGLAMPK 750
443 CKLQNLK 449
75: CQVKTLR 757

RESULT 7
ID_PVA5_RAT STANDARD; PRT: 354 AA.
AC C63035;
DT 28-FEB-2003 (Rel. 42, Last sequence update;
DT 28-FEB-2003 (Rel. 42, Last annotation update;
DE PVRIN-containing APAFI like protein 5 like (Anaktensin II)/vasopressin
receptor;
UN PYPAF5 OR AVR.
OR Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI TaxID:10116;
PP CONCEPTUAL TRANSLATION OF 1-343.
RA Rinz U...
RL Unpublished observations (FEB-2003).
RN 12.
RP SEQUENCE OF 144-854 FROM N.A., MUTAGENESIS OF SER 193, AND FUNCTION.
RX TISSUE=Kidney;
PC MEDLINE=96071640; PubMed=7459366;
RA Ruiz-Capazo N., Akimoto K., Herrera V.L.M.;
RT Identification of a novel dual aspartensin II/vasopressin receptor on
the basis of molecular recognition theory;
RL Rat. Med. 1:1074-1082(1995).
RN 13.
RP VARIANTS SALT-SENSITIVE HYPERTENSION SER 492 AND ARG 536.
EX PHELINE219K185; PubMed:1984503;
RA Ruiz-Capazo N., Lopez L.V., Herrera V.L.M.;
RT The dual AngII/AVP receptor gene N19S/184R variant exhibits
sodium-induced dysfunction and coexpressed with salt sensitive
hypertension in the Dahl salt-sensitive hypertensive rat model;
RL Mol. Med. 8:24-32(2002).
RN 14.
RP IDENTIFICATION OF XANMALIAN GTP-BINDING PYPAF5.
RA Albright M., Domingues F.S., Schreiber S., Lemmon T.;
RT Identification of mammalian orthologs associated with GTPAF5 with distinct
functional roles;
RL FEBS Lett. 538:173-177(2003).
CC 1. FUNCTION: May mediate activation of GASP1 via ASC and promote
activation of NF-kappa-B (by similarity). Anaktensin II and
vasopressin binding protein. May stimulate cAMP accumulation.
CC 2. SUBUNIT: Binds to ASC with its CAPIN domain (by similarity).
CC 3. SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC 4. TISSUE SPECIFICITY: Highly expressed in kidney. Detected at low
levels in all tissues tested.
CC 5. DISEASE: Defects in PYPAF5 may be a cause of salt sensitive
hypertension.
CC 6. SIMILARITY: Contains 1 CAPIN domain.
CC 7. SIMILARITY: Contains 1 NACT domain.
CC 8. CAUTION: The N-terminus was extended using ESTs and genomic
sequences, in analogy to ortholog sequences.
CC 9. This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.
EMBL: M85183; AAA33623.1; ALT_INIT.
DR InterPro: IPR001411; LRR_RNinh
DR Pfam: PF00560; LRR_2.
DR Pfam: PF02758; PAAD-DARIN_1.
DR PROSITE: PS50624; DARIN_1.
DR PROSITE: PS50837; NACT_1.
KK ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 133 CAPIN.
FT DOMAIN 144 484 NACT.
FT DOMAIN 544 634 ASP/GLY RICH
FT DOMAIN 635 662 POLY-LYS
FT REPEAT 433 458 LRR 1.
FT REPEAT 612 633 LRR 2.
FT REPEAT 722 745 LRR 3.
FT REPEAT 812 836 LRR 4.
FT NP_BIND 174 181 ATP (POTENTIAL).
FT VARIANT 492 492 N -> S IN SALT-SENSITIVE HYPERTENSION.
FT VARIANT 536 536 C -> R IN SALT-SENSITIVE HYPERTENSION.
FT MUTAGEN 766 766 E -> K; ABOLISHES ANGIOTENSIN II BINDING.
SQ SEQUENCE 354 AA; 95290 MW; DHRG922D7B3734 CRC64.
Query Match: 27.6%; Score 695; DB 1; Length 854;
Best local similarity: 32.1%; Pred. No. 3.9e 47;
Matches 170; Conservative 92; Mismatches 114; Gaps 13;
QY 9 SMOGLFENPSPAPQGLIRVPERLLFIIDGPDLPKSPHDGPGACLQWELKRPTEL 68
DB 220 SRAHCEGCPQPCAPVPRILAQHRLFLIUGADELPTAAPATPC----PCPEA 273
QY 69 -- LENSIPKKLAFELSSLTTPFALEXHRLLEPRHVELGFSEAEKREYFYKY 123
DB 274 TSGLRVLSGLSQELSPASALLVTSRNATLGLQGLSPCAEVRGFSQDKKKYFKE 333
QY 124 ENMAEAGQVNVVRDNEPFTKVFELVWVVTCTLOQLEGGLLQCTSTTTAVNL 183
DB 334 FQDERKAEAYRFVRENETVYALCFVFPVCMVTCVTLQQKELGRDLSTRTSTTSVY 393
QY 184 YLLSLVQPKFGAPRCQPPNCRGLCSLAAGLWCKILFEEODLRKKHGLQEDVSA 239
DB 394 FITSKFKAGSTNPPVQC--GELRMCHLAESILKQGFSEKDLERLKQSCVCTMFL 451
QY 240 --- FUNKN--- FQKJNCE RSFL 257
DB 452 SKRELEPTELVTVYTCFIDQSPQFELALNSYLDAEGAPGNSAGSVQMLRSALRGHL 511
QY 458 AATSPRELQELNEFESHLEKSKCKKVAAPHIKWALQWICSKAQSGSTL----- 307
DB 512 AATIRFPLGLNSTERIHCQNHFGVVPQPVQDQLRWVQGSQPKVATVGAKKDELKC 571
QY 308 --- QQGLEFFSKLYETQEEFQQALSHFQVIVWSNTASKMEHMYVS 351
DB 572 SEAEHREHEHEHEHEKAFGLGLLYCLVETCEDQFVRQALSSLPENYLERVRLTRVDLEV 631
QY 352 SPCLKRCRAQVHLKATYSGDGPAP-----CSAGAH---TLVQLPERIV 397
DB 632 SYCVQCPDQALRVSCGLVAAREKKKKSLVKEKSGSOSTKKQPPVSMRELCEMT 691
QY 398 --- LLDAYSEHMAALCTNPLELSCYRNALGSRGVKYLCCGGLRHN 442
DB 692 TQCGHLSLTLSHCKLPDAVCRDGLFAKVAFALESLVSLTETGFLCEGLAMPK 751
QY 443 CKLQNLK 449
DB 752 CQVKTLR 758
RESULT 8
NAL4_HUMAN STANDARD; PRT: 994 AA.
AC Q96MN2; Q96AV6;

```
CC CC      :isold-Q96MN2-2; Sequence:VSP_003917;
CC Name=3;
CC      :isold-Q96MN2-3; Sequence:VSP_003916;
CC -! SIMILARITY: Contains 1 DAPIN domain.
CC -! SIMILARITY: Contains 1 NACTH domain.
CC -! SIMILARITY: Contains a leucine-rich LRR; repeats
CC -! CAUTION: A stop codon in Ref.4 was read through in position 4; to
CC extend the sequence.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement [see http://www.isb-sib.ch/attribution/
CC or send an email to licensais@sib.ch].
CC
CC EMBL; AF442489; AAL35293.1;
CC EMBL; AY372792; AAL68396.1;
CC EMBL; AY379747; AAL91704.1;
CC EMBL; AF482706; AAL86722.1; ALT_INIT.
CC EMBL; AK566893; BAB2254.1;
CC EMBL; RC316443; AAI16443.1;
CC InterPro: IPR007091; LRR_RNase.
CC InterPro: IPR007111; NACTH_NTPase.
CC InterPro: IPR004023; PAAD_DAPIN_dom.
CC Pfam: PF02759; PAAD_DAPIN_1.
CC PROSITE: PSSC804; DAPIN.
CC PROSITE: PSS0837; NACTH; 1.
CC Repeat: Leucine-rich repeat; Alternative splicing.
CC DOAIN 1 94 DAPIN.
CC FT          149 472 NACTH.
CC FT REPEAT   637 662 LRR 1.
CC FT REPEAT   723 745 LRR 2.
CC FT REPEAT   750 773 LRR 3.
CC FT REPEAT   780 797 LRR 4.
CC FT REPEAT   806 833 LRR 5.
CC FT REPEAT   853 886 LRR 6.
CC FT REPEAT   920 943 LRR 7.
CC FT REPEAT   949 972 LRR 8.
CC FT VARSPLIC ; 93
CC FT          MASPFPSQGLMWYLEELKEEFKFKKH;KMTLO;LSKC
CC FT IPTVEVKASREELANLLIKIVRCQAWNITIRIPQDPKP
CC FT QG-MKWVPEPT->MQECCLLVFSP;ALTSS [in
CC Inform 3].
CC FT          /rid=vsp_003916.
CC FT Missing in isoform 2).
CC FT          /rid=vsp_003917.
CC FT          L->P [in REF. 4].
CC FT Y>H [in REF. 4].
CC FT I->I [in REF. 4].
CC FT W>R [in REF. 4].
CC FT I>S [in REF. 4].
CC FT Y>C [in REF. 4].
CC FT G>A [in REF. 4].
CC FT L>P [in REF. 4].
CC SEQUENCE 994 AA: 1:3444 MW: iCECTCG69B1839BC9 CRC64;
CC
CC Query March 27'53; Score 652.5; DB 1; Length 994;
CC Best Local Similarity 3114; Pred. No. 78e-470;
CC Matches 153; Conservative 98; Mismatches 170; Indels 71; Gaps 7;
CC
CC QY 9 SMOCLIFSCWPEPSARLQSLRVPERLLFIIDGFDELKPSFDPGQPWCMLCWEEKRPTEL 68
CC Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 200 SJALLISREPDPNAPIETVISQPELLFVIDSFEELOGGNEPDSDLCDLMEKRFVQV 259
CC
CC QY 69 LENSILRKLLPELSLSITTRPTALEKRHLRLHPRHVEILGFSEAEERKYFYKFNAAE 128
CC Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 260 LSSSLPLKKWLPEASLSIAIKVPCKELRDQVTISEIVQPPGFNESRLVYFCFFCKDKP 319
CC
CC QY 129 QAGGVNYVDNEPLTYKCFVPLVNWVVCTTLQQQLGGCLLRQTSRTTTAVMYCYLLSI 188
CC Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 300 RAMEARNLVEESQSSEIQILLWMILCTSLKQPMQGMKDLAJTCSTTSVSYVSFVENL 379
CC
CC QY 189 YAKKIWA PFLCPPPNQSGLSHAADGLPWOKLFIEQDLRRKHGLDGEVSAFLNNINF 246
```



```

RA      Lora J.M., Geddes S.C., Atkinson N.J., DiStefano P.S., Reifin J.J.  
RT      "PYPAP7, a novel pYRN-containing Apaf-1-like protein that regulates  
RT      activation of NF-kappa B and caspase 1-dependent cytokine  
RT      processing." J. Biol. Chem. 277 29874 29880(2002).  
RJ      J. Biol. Chem. 277 29874 29880(2002).  
CC      -!- SIMILARITY: Contains 1 DAPIIN domain.  
CC      -!- SIMILARITY: Contains 1 DAPIIN domain.  
CC      -!- SIMILARITY: Contains 9 Leucine-rich (LRR) repeats.  
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation,  
CC      the European Bioinformatics Institute. There are no restrictions on its  
CC      use by non-profit institutions as long as its content is in no way  
CC      modified and this statement is not removed. Usage by and for commercial  
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement/  
CC      or send an email to license@sib.ch).  
DR      EMBL: AF464765; AAL69361.1;  
DR      InterPro: IPR007091; LRR_NRP_nrh.  
DR      InterPro: IPR007111; NACHT_NTPase.  
DR      InterPro: IPR004020; PAAD_DAPIIN_dom.  
DR      Pfam: PF027366; PAAD_DAPIIN_1.  
DR      PROSITE: PSS0H24; DAPIIN_1.  
DR      PROSITE: PSS0E17; NACHT_1.  
KW      ATP-binding; Leucine-rich repeat; Repeat.  
FT      DOMAIN : 93 DAPIIN.  
FT      DOMAIN : 172 49; NACHT.  
FT      REPEAT : 614 638 LRR 1.  
FT      REPEAT : 674 697 LRR 2.  
FT      REPEAT : 760 784 LRR 3.  
FT      REPEAT : 798 810 LRR 4.  
FT      REPEAT : 817 847 LRR 5.  
FT      REPEAT : 845 869 LRR 6.  
FT      REPEAT : 874 907 LRR 7.  
FT      REPEAT : 903 928 LRR 8.  
FT      REPEAT : 933 957 LRR 9.  
FT      REPEAT : 174 195 LRR 10.  
FT      N_BIND : 980 984 ATP (POTENTIAL).  
SQ      SEQUENCE 980 AA; 111836 MW; 822AF2FD413B303D CRC64;  
  
Query Match 21.1%; Score 561.5; DB 1; Length 380;  
Best Local Similarity 25.1%; Ref No. 2 aa-39;  
Matches 169; Conservative 8; Mismatches 129; Indels 24; Gaps 14;  
  
CY      R CSMGRLLEFCWPSRPLQESRYPERGLFIIDGFDELKPSFHDPOGW...CLCWEEK 63  
CB      LRI CRPAALLIKRWPELOLLIRISIAQAQRIFFWDSGLDLKV...PCALICDICGDWEKK 276  
QY      64 RPTEILLNLSRKELIPSLSTLTPTALEKHLLEHPRHVEIIGFSASKEYFYKY 123  
DB      277 KPVVIVLSSLKGVMPRAAILVTTPRALRLQLLAQQPIYRVVGFLSDRRATFLRH 336  
QY      124 FHNARAGGVNYVRNDEPLFTMCVPVLKCVWVCCTQQOLEGGGLRQTRTTAVYM 183  
DB      337 FSPEDQAVRAFELMRGNALFQGSGAPAVWIVCTTLKLQMEKGDPVPCLTRTGFLR 336  
QY      164 VLSLRCPGPCARLAFPPNQRLGSLAAGLNKNKILSEQCRLRXHLDGEDVSFAFLM 243  
DB      397 FLCSRF--FGGAQGRPA--RFLSMAAGCWAQMSVFHREDLERJGVGESDRFLFG 451  
QY      244 NIFORD-----INCERSFLAL----- 259  
DB      452 QILRLSVSKGCYSF-HLSPCOPLTALYALEKEGEDRDGHAWDIGVKLI-SGEERLK 511  
QY      260 -----TSRFLPGLLNEITRSULEKSCKWKVSPHIKMOLLQHWIQSAQDGST-QQSLE 313  
DB      512 NPDIQQVHFHFLGLAKREKAKELEATTGGRMSPD;KOELLQ-CKAHLCHANKPLSV-DLKE 570  
QY      314 FFQSYVEIEDEEFIOGALSHQGVIVYSNASKMEHVSVSFCLKRCPSAGVHLHYGA--- 369  
CB      571 VLGTLESSEEEELAKVVVAFFKRISI-HLTNTSEVPHCSFSLKHGOOLKLS-QVAKGVF 629  
QY      370 -----TVSAD 374

```


"A mouse gene encoding an oocyte antigen associated with autoimmune premature ovarian failure";
 Endocrinology 140:3720-3726(1999).

SEQUENCE FROM N.A. AND VARIANTS ASP-7; LYS 42; LYS-87; PRO-256;
 GLN-977 AND 1003-ASN ASN-1004.

STRAINS=29/SV;
 MEDLINE=20222768; PubMed=12751091;

Tong Z.-B., Nelson L.M., Dean C.
 "Water encodes a maternal protein in mice with a leucine-rich repeat domain homologous to porcine ribonuclease inhibitor";
 Yarn. Genome 11:281-287(2000).

FUNCTION:
 ME21NF=20517328; PubMed=11062459;

Tong Z.-B., Gold L., Pfeiffer K.E., Dorward H., Lee E., Hoady C.A.,
 Dean J., Nelson L.M.;
 "Water, a maternal effect gene required for early embryonic development in mice";
 Nat. Genet. 26:267-268(2000).

FUNCTION: Necessary for embryonic development beyond the 2-cell stage.

1- SUBCELLULAR LOCATION: Cytoplasmic.
 TISSUE SPECIFICITY: Oocyte-specific.

2- SIMILARITY: Contains 1 NACT domain.
 3- SIMILARITY: Contains 13 leucine rich (LRR) repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed (stage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/submit/announce/> or send an email to license@ebi.ac.uk).

EMB1: AF074018; AAD5762.1;
 EMB2: AF143573; AAF64393.1;
 EMB3: AF143559; AAF64393.1; JOINED.
 EMB4: AF143560; AAF64393.1; JOINED.
 EMB5: AF143561; AAF64393.1; JOINED.
 EMB6: AF143562; AAF64393.1; JOINED.
 EMB7: AF143563; AAF64393.1; JOINED.
 EMB8: AF143564; AAF64393.1; JOINED.
 EMB9: AF143565; AAF64393.1; JOINED.
 EMB10: AF143566; AAF64393.1; JOINED.
 EMB11: AF143567; AAF64393.1; JOINED.
 EMB12: AF143568; AAF64393.1; JOINED.
 EMB13: AF143569; AAF64393.1; JOINED.
 EMB14: AF143570; AAF64393.1; JOINED.
 EMB15: AF143571; AAF64393.1; JOINED.
 EMB16: AF143572; AAF64393.1; JOINED.
 PIR: A59000; A59000.
 HSSP: P10775; 2BNH.
 MOP: YG11345.93; Mater.
 GO: GO:0005829; Cytoosol; IDA.
 GO: GO:0009887; P:organoogenesis; IPI.
 InterPro: IPR001621; LRR.
 InterPro: IPR007391; LRR RNinh.
 InterPro: IPR021111; NACT_NTPase.
 Pfam: PF00560; LRR_3.
 PROSITE: PSS0637; NACT; 1.
 ATP-binding: Leucine-rich repeat; Repeat; Poly-nip-sm.
 DOXAIN 1 128 5 X APERXINATE TADDEM REPEATS.
 REPEAT 1 23 1 (INCOMPLETE).
 REPEAT 24 49 2.
 REPEAT 50 75 3.
 REPEAT 76 102 4.
 REPEAT 103 128 5.
 CORAIN 191 513 NACT.
 REPEAT 540 664 IPI.
 REPEAT 691 714 LRR 2.
 REPEAT 737 762 LRR 3.
 REPEAT 776 799 LRR 4.

FT REPEAT RC4 832 LRR 5.
 FT REPEAT 834 856 LRR 6.
 FT REPEAT 861 884 LRR 7.
 FT REPEAT 890 913 LRR 8.
 FT REPEAT 918 941 LRR 9.
 FT REPEAT 950 974 LRR 10.
 FT REPEAT 978 1003 LRR 11.
 FT REPEAT 1025 1059 LRR 12.
 FT NP BIND 1032 1059 LRR 13.
 FT VARIANT 197 204 ATP POTENTIAL.
 FT VARIANT 42 42 E -> D 12N STRAIN 129/SV.
 FT VARIANT 87 87 R -> K 12N STRAIN 129/SV.
 FT VARIANT 256 256 L -> P 12N STRAIN 129/SV.
 FT VARIANT 977 977 H -> Q 12N STRAIN 129/SV.
 FT VARIANT 1003 1004 SS -> NN 12N STRAIN 129/SV.
 SQ SEQUENCE 1111 AA; 125502 MW; 4016A5D67A1CC1F4 CRC64;
 Query Match 22.4%; Score 545; DB 1; Length 1111;
 Best local similarity 29.1%; Pred. No. 6.9e-37;
 Matches 143; Conservative 85; Mismatches 207; Indels 56; Gaps 6;
 QY 9 SMQLIFSCWPSAPLQELIVPERLFLICGFDELKPSFDPOOPWCLCHEKRPTEL 69
 DB 241 SLAQLIAKCPDSNOLVTKIMSQPELLFVLSGLDEMDSVLQHDQMT-SRCWKQSQPIVI 300
 QY 69 LNSLIRKKLDELSELITTRPTALEKJHRLLEHPRHVEILGFSEAEKKEYKVFHNAE 128
 DB 301 LKYSILRKALQCSFELITTRNTGLEKLSMYYVSPYILVESLSASRRSCVLIENISNES 160
 QY 129 QAGQVENVYRDEPFPKCFVPLVQWVTCIQQLQEGGLLRQTSRTTAYV MIYL 185
 DB 361 LRIQVHSLIENHQFPDQCAPSVCSLNCALQKLLKRCCTLPCCNLGLDYATLVHQ 420
 QY 186 LSIQMKKAPRLQPTNORSCSLAAGCQKQKLFSEQDLRKHGLGCVSAFLNNNI 245
 DB 421 LTRKPSQSAISQERQITLVGLQMAAGVMTMRSFYDDDLKNYSLKESLALAFKNI 480
 QY 246 FQKQNCERS-----FLAL----- 259
 DB 481 LLOVGHNSQCYVFHSHLSQDFEALYVLEGEENQHFCEINQPSIYEVKRTDTRL 540
 QY 260 LSKPLGLINEFTRSLKSLQKVKVSPH:KYVLLQWISKAQSGSTLQGSSEFFSC 317
 DB 541 LGVKEPRLGLMKDILKLELVILFVTVPTVEKQLHWVSLIAQCVNGTSPMDTLDAFYC 600
 QY 318 LVELGEEELQCALSHFOVLVSNASKWEHYVSPFCIKRCSAQVLIHLYGATYSACGED 377
 DB 401 LFESQDERFWGALKRFPQEWILLNQKMKVSSYCYLKHQCNLKAIRVETEDLLSVDMT 659
 QY 378 RARCSAGAHILVQIPERTVLLDAYSEHLAAALCTNPDLIELSLYRNALSGVGLCOG 437
 DB 660 LRLCP---VVIVQETQCPFLMEWGNFCSVLGSLRNKELDLGDSILSORAMKLLCLE 715
 QY 436 LRHPNCKCNL 448
 DB 716 LRNCSQRIQK 726
 RESULT 13
 PYA6_HUMAN STANDARD; PRT: 1033 AA.
 AC PS9045; Q8NSF5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, last sequence update)
 DT 15-SEP-2003 (Rel. 42, last annotation update)
 DE PYRIN-containing APAF1-like protein 6.
 GN PYPAF6.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Vertebrata; Euteleostomi;
 CC Mammalia; Eumetazoa; Primates; Catarrhini; Hominoidea; Homo.
 CX NCB:TaxID:9606;
 RN [1]

SEQUENCE FROM N.A. (ISOFORM 1);
 MEDLINE=22162427; PubMed=12019269;
 Wang L., Manji G.A., Grenier J.V., Al-Garawi A., Veritia S.,
 Ibra J.M., Geddes B.J., Brisken C., DiStefano P.S., Bertin J.,
 "Pvpa2", a novel p39IN-containing Apaf-1 like protein that regulates
 activation of NF-kappa B and caspase-1-dependent cytokine
 processing.;
 J. Biol. Chem. 277:29874-29885 (2002).
 [2]
 SEQUENCE FROM N.A. (ISOFORM 1);
 MEDLINE=22162427; PubMed=12019269;
 Strausberg R.L., Feingold E.A., Grouse L.H., Bergo M.O.,
 Klausner R.D., Collins F.S., Wagner L., Shenman C.V., Schuler G.E.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Vax S., Wang J., Hsieh F.,
 Chacklenko L., Narasimha K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ustin B., Tomiyaki S., Carninci P., Prange C.,
 Rata S.S., Lequellier N.A., Peters G.J., Aurora R.D., Millan S.C.,
 Bonak S.A., McEwan P.J., McKernan K., Valle D.A., Gharatne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., May L.J., Bulvik S.N.,
 Villalón D.K., Kuzny D.W., Sodergren E., Lu X., Gibbs P.A.,
 Bailey A., Helton E., Kottmann M., Vadan A., Rodriguez S., Sanchez A.,
 Whitting M., Madan A., Young A.C., Stevchenko Y., Burdick G.G.,
 Sakantey R.W., Touchman J.W., Green E.B., Ileson M.C.,
 Rodriguez A.C., Grimwood J., Schatz G., Myers R.W.,
 Butlerfield V.S.N., Kryzanski V.J., Skalska J., Stallus D.E.,
 Schreier A., Schein G.E., Jones S.J.N., Maira M.A.,
 "Generation and initial analysis of more than 15,000 full length
 human and mouse cDNA sequences.;
 Proc. Natl. Acad. Sci. U.S.A. 99:16499-16502 (2002).
 [3]
 SEQUENCE OF 365-1033 FROM N.A. (ISOFORM 2);
 TISSUE=Glial tumor;
 Kawakita B., Suiyara A., Takemoto M., Suiyara T., Irie R.,
 Chiku T., Sato H., Ota T., Kakamatsu A., Ishii S., Yamaguchi C.,
 Isoro Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 Yamashita H., Matsuo K., Nakamura Y., Sekine Y., Kikuchi H., Kanda K.,
 Magatsuma M., Murakawa K., Kanohori K., Takahashi-Fujii A., Oshima A.,
 Suzuki Y., Sugano S., Nagahara K., Masuko Y., Nagai K., Isogai T.,
 "NEO human cDNA sequencing project.
 Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 [4]
 ALTERNATIVE PRODUCTS:
 Event: Alternative splicing; Name: Isoform 1;
 Name: 1;
 IsoId: P59045-1; Sequence Display:
 Name: 2;
 IsoId: P59045-2; Sequence Display:
 Note: No experimental confirmation available.
 [5]
 SIMILARITY: Contains 1 DAPIN domain.
 [6]
 SIMILARITY: Contains 6 leucine-rich repeat repeats
 [7]
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL database
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed, usage by and for commercial
 entities requires a license agreement. See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch.
 [8]
 EMBL: AY095145; AAM14632.1;
 EMBL: BC034730; AAH34730.1;
 EMBL: AK090621; BAC03490.1; ALT INIT
 InterPro: IPR037091; LRR_RN: h
 InterPro: IPR033590; LRR_RN: h
 InterPro: IPR037111; NACHT_NTFase
 InterPro: IPR034020; PAAD_DAPIN_1
 Pfam: PF00560; LRR; 1
 Pfam: PF02758; PAAD_DAPIN; 1
 SMART: SM00368; LRR_R1; 9
 PROSITE: PS50824; DAPIN; 1

PROSITE: PS50817; NACHT; 1
 ATP-binding; leucine-rich repeat; Repeat; Alternative splicing.
 KW DOMAIN 1 91
 DAPIN
 FT REPEAT 149 473
 NACHT
 FT REPEAT 549 611
 LRR 1
 FT REPEAT 632 658
 LRR 2
 FT REPEAT 745 788
 LRR 3
 FT REPEAT 802 827
 LRR 4
 FT REPEAT 849 882
 LRR 5
 FT REPEAT 919 944
 LRR 6
 FT NP_BIND 153 158
 ATP_BIND
 FT VARSPIN 614 688
 SOMKSLVYKREKSLFYFMESRELHIFNDLNGISERILS
 KALEHSGKRLK -> R (in isoform 2)
 /FTID-VSP 00768
 T 3 1 IN REF. 31
 FT CONFLICT 788 788
 T 3 1 IN REF. 31
 SQ SEQUENCE 1031 AA; 117994 MW; A25880485FBA49AC CRC64;
 Query Match 18.73; Score 428.51; DB 1; Length 1033.
 Best Local Similarity 26.23; Pred. No. 5.6e-28;
 Matches 111; Conservative 85; Mismatches 211; Indels 71; Gaps 12;
 QY 1 MKGATEKSKVQQLIFSCWPEPSAPQELIRVPERLLEFIDGDELKPSFHPDQGPWCLOW 60
 DB 19: IKQ-KTSSSLASLAKDKWGGGAPADLSLKKKLLFLELDLNREFLNVESALCSNS 249
 QY 61 EKKPTTEALLNSLRKKKSLPELSSLTTPRTALEKHLLEHPRHVEILGISEAKRYEYF 120
 DB 250 TOKVPIPVLLVSLKKKMAPGWFTLSRPTRGKVKTKLKEVDCQTTLSLNGKREIYF 309
 QY 121 YAYHNAAAGQGVNRYVNEPELFTYCFVPLVGVVVCTCLOOLEGGGLPQTSSTTAV 180
 DB 310 NFFKQQRASAAQLVHEDEILVGLCRVALLCWITCTVLKQNDKGRQFQCCQFTDL 369
 QY 181 YMYLLISLQPKPAPRLQFPNQ-----RGLCSLAADGLMNKILFEEQPLRKHSLDG 234
 DB 370 HAFLADALTSAGL-----TANQVHGLGLKPLLLAAGGLFSLTINSGEDLRGVGTE 424
 QY 235 EVNSRFLKNKIF-----QKD-----INCERSFLAT-----SR----- 262
 DB 425 ADVSVLIQAANILLPSNTHRDYKFIHLNQVQFCTATFLAVAVNYLLPSGSREYKREK 484
 QY 263 -----FLRLNNEETRSLEKSLCKWKVFKMDLQW-----LQSKASDQSTLQO 309
 DB 486 YSDPKGVFTFEGLLNANERKLETSPGYQCP---MVDSFKWYSGYMKHLDPEKILTH 541
 QY 310 GRLEPFTVYECSEEFICQALSHFOVIVYSNTAKMHWSSFCGLKCRSAQVHLHYGA 369
 DB 542 HMEFLYCVYENRESEFFVKTVDALWEVTVYQSKOMVSYLYCLDYCHLRTLKLSVQ 599
 QY 370 TVSAQVEDPARGASANTLVOLPERTVLDAYSSEHLAALCTNPNLIELSYRNALGSR 429
 DB 600 RLPQKPELRPTASQMSLV-----VYREICSLFYTMESLRELHIFDNDLNGI 648
 QY 432 GYKLLQCGRHNPCKLQNR 449
 DB 649 SERLSKALEHSSCKRLTK 668
 RESULT 14
 CARE_HUMAN STANDARD; PRT: 1040 AA.
 AC Q9HC29; Q96RH5; Q96RH6; Q96RH9;
 DT 28-FEB-2003 (Rel. 41, Created;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Caspase recruitment domain protein 15 (Nod2 protein); (Inflammatory
 bowel disease protein 1);
 GN CARD15 OR NOD2 OR IBD1
 OS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]

GenCore version 5.1.6
Copyright (c) 1993 - 2001 GenPhen Ltd.

OK Protein - protein search, using sw model

Run on: October 31, 2003, 06:54:44 Search time 114 Seconds
without alignment!
1016.365 Million cell updates/sec

Title: US-10-028 374 2

Perfect score: 2377

Sequence: 1 MNSATECSQDLIFSCWPE.....GVKLLGVGLSHPRKCNLR 449

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 256052604 residues

Total number of hits satisfying chosen parameters: 810525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

- 1: sp_archaea:
- 2: sp_bacteria:
- 3: sp_fungi:
- 4: sp_human:
- 5: sp_invertebrate:
- 6: sp_mammal:
- 7: sp_rhiz:
- 8: sp_organellae:
- 9: sp_phage:
- 10: sp_plant:
- 11: sp_rodent:
- 12: sp_virus:
- 13: sp_vertebrate:
- 14: sp_unclassified:
- 15: sp_vivus:
- 16: sp_bacteriap:
- 17: sp_archaeap:

Pred. No. is the number of results predicted by charge to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	2334	98.2	892	4	Q8NEU4	Q8NEU4 homo sapien
2	1145	48.2	1035	11	Q8R4B6	Q8R4B6 mus musculu
3	594	25.0	713	6	Q95127	Q95127 macaca fasc
4	563	23.7	982	11	Q8R440	Q8R440 mus musculu
5	549.5	23.2	846	4	Q8IX10	Q8IX10 homo sapien
6	519	21.8	963	11	Q8CCJ9	Q8CCJ9 mus musculu
7	494	20.8	748	11	Q99MWC	Q99MWC mus musculu
8	473.5	19.9	673	11	Q8CKN1	Q8CKN1 mus musculu
9	443	18.6	825	11	Q8C6M5	Q8C6M5 mus musculu
10	393.5	16.6	657	11	Q9EP37	Q9EP37 mus musculu
11	261	11.0	953	4	Q8WF5	Q8WF5 homo sapien
12	253	10.6	1020	11	Q8K120	Q8K120 mus musculu
13	246.5	10.4	778	4	Q8NF48	Q8NF48 homo sapien
14	233	9.8	736	11	Q8BUT6	Q8BUT6 mus musculu
15	233	9.8	953	11	Q8B8R0	Q8B8R0 mus musculu
16	223	9.4	660	4	Q8NF65	Q8NF65 homo sapien

17	212.5	8.9	519	11	Q8C243	Q8C249 mus musculu
18	212.5	8.9	864	7	Q29675	Q29675 homo sapien
19	212.5	8.9	932	4	Q96K14	Q96K14 homo sapien
20	280	8.4	682	4	Q96D51	Q96D51 homo sapien
21	176	7.4	1052	7	Q9GCD8	Q9GCD8 rattus norv
22	176	7.4	1073	7	Q9GCD9	Q9GCD9 rattus norv
23	176	7.4	1153	7	Q9GCE0	Q9GCE0 rattus norv
24	174	7.3	977	7	Q8HW99	Q8HW99 mus musculu
25	174	7.3	1155	7	Q9TTP1	Q9TTP1 mus musculu
26	165.5	7.0	516	4	Q8NRF5	Q8NRF5 homo sapien
27	165.5	7.0	1056	4	Q8NRF4	Q8NRF4 homo sapien
28	155	6.5	733	4	Q8NFE2	Q8NFE2 homo sapien
29	150.5	6.3	1064	5	Q94429	Q94429 ciona intes
30	118	5.8	1097	4	Q9H6Y0	Q9H6Y0 homo sapien
31	135.5	5.7	287	4	Q9H125	Q9H125 homo sapien
32	135	5.7	509	4	Q9H529	Q9H529 homo sapien
33	116.5	4.9	1075	5	Q94430	Q94430 ciona intes
34	113.5	4.9	393	11	Q8BWC2	Q8BWC2 mus musculu
35	111	4.7	269	11	Q92458	Q92458 mus musculu
36	110	4.6	456	11	Q91V17	Q91V17 mus musculu
37	109	4.6	393	5	Q81FTR	Q81FTR marthasteri
38	108	4.5	461	6	Q8H2P4	Q8H2P9 pan troglod
39	107	4.3	447	4	Q96FD7	Q96FD7 homo sapien
40	107	4.3	461	4	Q9HC80	Q9HC90 homo sapien
41	107	4.3	461	4	Q81ZK9	Q81ZK9 homo sapien
42	107	4.3	2530	12	Q8BEN2	Q8BEN2 crotopouche v
43	106.5	4.3	390	13	Q98T12	Q98T12 xenopus lae
44	106.5	4.3	1160	4	Q81Z24	Q81Z24 homo sapien
45	106	4.3	456	11	Q924P4	Q924P4 mus musculu

ALIGNMENTS

RESULT 1

Q8NEU4	PRELIMINARY;	EPT;	892 AA.
ID	Q8NEU4		
AC	Q8NEU4		
DT	01-OCT-2002 (TREMBL)	22, Created:	
DT	01-OCT-2002 (TREMBL)	22, Last sequence update:	
DT	01-MAR-2003 (TREMBL)	23, Last annotation update:	
DE	Monarch:1 splice form IV.		
OS	Homo sapiens (Human)		
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
CX	NCBI TaxID:9606;		
RM	11;		
RP	SEQUENCE FROM N.A.		
RA	Williams K.L., Lischoff M.W., Tang J.P.Y.,		
RL	Submitted (MAY 2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; A116207; AAK75145.1;		
DR	InterPro; IPR007091; LRR_RNrich.		
DR	InterPro; IPR007111; NACHT_NTPase.		
DR	InterPro; IPR004020; PAAD_DAPIN_dom.		
DR	Pfam; PF02758; PAAD_DAPIN_1.		
DR	PROSITE; PSS0824; DAPIN; 1.		
DR	PROSITE; PSS0503; LRR_R1; 1.		
DR	PROSITE; PSS0817; NACHT; 1.		
SQ	SEQUENCE 892 AA; 101733 MW; AE703D8DF341C2AC CRC64;		

Query Match: 98.2%; Score 2334; DB 4; Length 892;
Best Local Similarity 90.7%; Pred. No. 26-218;
Matches 449; Conservative 0; Mismatches 0; Indels 46; Gaps 2;

QY	1	MNSATECSQDLIFSCWPEAPQLIRVPERLLFIIDGDELKPSFHDQGGKCLCW 60
DB	255	MNSATECSQDLIFSCWPEAPQLIRVPERLLFIIDGDELKPSFHDQGGKCLCW 314
QY	61	EKKRPTTEMLNSLRKKLPELSCTTRPTALEKHRLLEHPRHVEIIGFSEARKEYF 120
DB	315	EKKRPTTEMLNSLRKKLPELSCTTRPTALEKHRLLEHPRHVEIIGFSEARKEYF 374
QY	121	YKFNHNASGAGQVFNVRNNEPLFNCVFPLVCWVVVTCLOQQJEGGGLRLRTSRTTAV 180

```

DB 375 YKFEHNAEQGVVYVRNREPLSTWTFVGVVWVWVTCCLQGLGGSLRQTSRTTAV 434
CY 181 YVYVLLSNOCKPKAPRLQPPENQPGGLCSAAGDKWCKTLIFEPONRKSGLEDGDSYAF 243
DB 435 YVYVLLSNOCKPKAPRLQPPENQPGGLCSAAGDKWCKTLIFEPONRKSGLEDGDSYAF 494
CY 241 LNMKIFQKDNK 255
DB 495 LNMKIFQKDNK 554
CY 256 FLAITSREPLFOLLNEERSRHSKSLCKWVSPIKMDLLOKIOSKAGSGSTLQGSLEFF 315
DB 555 FLAITSREPLFOLLNEERSRHSKSLCKWVSPIKMDLLOKIOSKAGSGSTLQGSLEFF 614
CY 316 SOLVEIOEEEFIOQALSHFOVIVVSNIAASKMHVSVSFCILKPSAQVHLYGATYADG 375
DB 615 SOLVEIOEEEFIOQALSHFOVIVVSNIAASKMHVSVSFCILKPSAQVHLYGATYADG 674
CY 376 EDRARCAGAGHTLAVQLPRTVLDAYSEHLAALQFNINLILSIYRNALSGRYKCL 434
DB 675 EDRARCAGAGHTLAVQLPRTVLDAYSEHLAALQFNINLILSIYRNALSGRYKCL 734
CY 435 CQGLRHKNCKLQNR 449
DB 749 CQGLRHKNCKLQNR 749

```

RESULT 2
Q95LZ7 PRELIMINARY; ERT; 1033 AA.

AC Q95LZ7
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Mast cell maturation inducible protein 1.
UN MNO101.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN 1.
RP SEQUENCE FROM N.A.
RC STRAIN=BAE/CJ;
RA Kikuchi-Yanosita R., Kaya K., Taketomi Y., Taniguchi T., Saito T.,
EA Ishii S., Hasegawa Y., Suzuki N., Shimizu T., Hanaoka M., Kato T.,
SA Chano H., W., Sawada M., Inagaki N., Nagai H., Yabuchi M., Kudo I.,
RT Identification of inducible genes during in vitro maturation of mouse
bone marrow-derived mast cells to conventional mouse type-1 mast cells."
RJ Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486612; AAL90974.1;
DR InterPro; IPR01092; RHM basic
DR InterPro; IPR010611; LRR
DR InterPro; IPR007391; LRR RNish
DR InterPro; IPR003590; LRR RNish_sub
DR InterPro; IPR007111; NACHT_NTPase
DR InterPro; IPR034320; PAAD_CAPIN_dom
DR Pfam; PF02560; LRR; 2
DR Pfam; PF02758; PAAD_CAPIN; 1
DR SMART; SM00369; LRR_E; 1
DR PROSITE; PS50824; CAPIN; 1
DR PROSITE; PS50038; HCH; 1;
DR PROSITE; PS50503; LRR R1; 1;
DR PROSITE; PS50837; NACHT; 1;
SQ SEQUENCE 1033 AA; 118274 MW; 5324670366b11177 PC64;

Query Match 48.2%; Score 1146; ERT; Length 1033;
Best Local Similarity 45.2%; Pred. No. 2, 2e-102;
Matches 23; Conservative 83; Mismatches 127; Indels 70; Gaps 5;
CY 6 TCSNQCILFSCWPSAPLQCLWVPRRLFLIISGRLEIKPSHNDPQPCMLCWEKRRP 65
DB 164 TRSJAQLIVSWPDPNPFWKLNKKGRLLFNDGFTLQAFLEHICVCTDKCKAVR 323

```

CY 66 TELLNSLRKKKLCLELSULITRTALEK;HRLLEHPRHVEILGFSSEARKEYEYKYPH 125
DB 324 GQILLSSLRKKLLPKASULITRTVALEK;LHLLDHPHVEILGFSSEARKEYEYKYPS 383
CY 126 NAEQACVENVYRDNKHPK;FEVGVVWVWVTCCLQGLGGSLRQTSRTTAVMYLV 165
DB 384 NELQAREAPLQRENEVFTMGFIPLVGVWVTCGLQQMETGKSAQTSKTTTAVYVFE 443
CY 186 LSNQPKAPRLQPPENQPGGLCSAAGDKWCKTLIFEPONRKSGLEDGDSYAFNMNI 245
DB 444 SSLQSPGGLSEHLSFYVQLQGLCSAAGDKWCKTLIFEPONRKSGLEDGDSYAFNMNV 503
CY 246 FQKDNK 254
DB 504 FQKDNK 563
CY 255 SFALTSREPLFOLLNEERSRHSKSLCKWVSPIKMDLLOKIOSKAGSGSTLQ 309
DB 564 GKFEKYLIFVVRFLSGLVNQRTSYLKKLCKISQVRLELLKMWIEVKAKKLQWOP 623
CY 310 GSLEPFTIVSCEERTIOQALSHFOVIVVSNIAASKMHVSVSFCILKPSAQVHLYG 369
DB 624 SOLEFYCLVMEQEEVQVSAMHFKHILNLSRVDHVSVSFCILKCHRYKLTSLGF 581
CY 370 TYSAGCLNARCSAGHTLAVQLPRTVLDAYSEHLAALQFNINLILSIYRNALSG 417
DB 682 FNSPKREBEERPGGRPLDQGVPEPOTHVAVSSRLVNCCLTSSFCGFSLSLSTNPSIT 741
CY 418 ELSLYNALSGROVKLQGLRHKNCKLQNR 448
DB 742 ELSLYNALSGROVKLQGLRHKNCKLQNR 772

```

RESULT 3
Q95LZ7 PRELIMINARY; ERT; 713 AA.

AC Q95LZ7
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Hypothetical; 80.3 kDa protein.
OS Macaca fascicularis (Crab-eating macaque) (Cynomolgus monkey).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CX NCBI_TaxID=9544;
RN 1.
RP SEQUENCE FROM N.A.
RC TSSUE=Testis;
RA Hashimoto K., Osada N., Hiji M., Kusuda J., Tanuma K., Hirai M.,
EA Terao K., Sugano S.;
RT Isolation of novel full-length cDNA clones from macaque testis cDNA
libraries."
RJ Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071044; BAB66437.1;
DR InterPro; IPR007391; LRR RNish
DR InterPro; IPR007111; NACHT_NTPase
DR PROSITE; PS50503; LRR R1; 2
DR PROSITE; PS50837; NACHT; 1;
DR PROSITE; PS50837; NACHT; 1;
KW Hypothetical; Protein;
SQ SEQUENCE 713 AA; 80259 MW; 6F214C9B773F54DC CRC64;

Query Match 25.5%; Score 594; DB 6; Length 713;
Best Local Similarity 28.5%; Pred. No. 7e-49;
Matches 16; Conservative 81; Mismatches 174; Indels 154; Gaps 13;
CY 13 LYSWPEISAPQLRIPERLFIIDFDELRPSFHDPOGW-----CLCWEKRRPTL 68
DB 1 XVFEDWPEACDIFHLAQACKTLFVIGDFDELGA----PPGALIQICGWQEQKPPV 56
CY 69 LANSLSKRLIPESULITRTALEK;HRLLEHPRHVEILGFSSEARKEYEYKYPHNAE 128

```

DB      129  LAGSLLKRRKXKPKATLLVTRPRAKNIPLFAEQRTYIRVLFQEDBRAYFLPHGDEE 116
QY      129  QACQVENVYRDNEFLTMGFVLYVWLTCTQQLUESGHLAPGCGPTTAYVYVLSE 188
DE      127  QAMRAFLMRNNAALFQGSAPAVWVYVTTKLQWNEHDAFTLTSTHLFDFLOCS 175
QY      189  MOPKQAPRQPPNCRKGLSIANKGLKILKLEBQVPRKRLDQWVSANLKNIFCK 248
DB      176  -QPPQA---Q-IGALRAISLAAASUNKAMVCHBHELEKAVQVDSRDLFLDRIIRC 231
QY      249  D-----INCRSPAL----- 259
DE      232  DGWAKGYCFINLSPCFITALFYALKRGEHQRDRAKLIIVVKKLASREELKPPCL 291
QY      260  TSPFLGGLNBEHTRSHLEKSLQWYSFHKMELQWISKQKQSDSTLQQSGLEFFSC 318
DE      292  QACRFIFGLANKRVRKLELATEKWNSEFFKQELRCQVSEKSHFTNADREILCCI 349
QY      319  YEIOEEEFICQALSHFCVIVVSNIAKYMWNYSFCLARKS-----A 362
DE      350  YESQDEELKRYNAQFKELSLHMAVDLATSFCFARSCCHIFFADACRNICLALRGHK 407
QY      362  QVJLYGATYSADGEDR-----A-CSA----- 383
DE      408  TVTHIL-----TLOGTQCKXKGLALCVLRPRNINLEYGLVACSATQCAWZSLALHAN 462
QY      484  -----GAILTVQCP-----ENTVLDAYSEHLAAALCTNKL 416
DE      463  RSLMVNLSDNELDEGAKLTLTHHPCFQFLDENFELTEAKNGLAAQVVSRL 522
QY      417  ILSILVNALSGPVKLACQVLRHKNKLCN 448
DE      523  TRLCLAKNSKTKGVAFKFGHNYFFICAL 564

RESULT 4
CARV40
ID Q8H440 PRELIMINARY: 187; 562 AA.
AC Q8H440
DE 01-MAR-2003 (TXEMBLrel. 23, Created)
DT 01-MAR-2003 (TXEMBLrel. 23, Last sequence update)
DI 01-MAR-2003 (TXEMBLrel. 23, Last annotation update)
DE Similar to PAN2 protein.
CS Mus musculus (mouse).
CC Eukaryota; Metazoa; Chordata; Mammalia; Artiodactyla; Eulipotyphla;
CC Mammalia; Eutheria; Primates; Carnivora; Canidae; Mustelidae;
CX NCBI_TaxID:101090
KW [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Ovary;
RX MEDLINE=2354493; PubMed=12466451;
RA The RIKEN Genome Consortium.
RT "Analysis of the mouse transcriptome base set and functional annotation of
RC 52,370 full-length cDNAs."
RL Nature 423:561-573(2002).
TR EMBL; AK07843; BAC40241;
SQ SEQUENCE 982 AA; 112600 MW; 6E2CEFEACD5185E P 64.

Query Match 23.7%; Score 563; DP 11; Length 982;
Best Local Similarity 27.0%; Pred. No. 116 44;
Matches 143; Conservative 136; Mismatches 147; Indels 94; Gaps 9.

QY      9  SMOGLSCWPSPAPQLQELIRVFERLITIDRDELKSFQCPQPKLQWEEKRPTEL 68
DE      193  SGAELISRWPGDSAEELSKPEKALITLSEKNEELTPWSESLCONTEKQPVAV 258
QY      69  LKSLIRKLLPELSLITITRTALEKRLHLSHPHREVLTPSPAERKRYFYFYFNA 127
DE      259  LSSLLRKNLPESSLEISATFESKKNRYLEYTHNFTFQKLNKNTKMSPHRLFQPR 317
QY      128  EQACQVENVYRCNEPFLTMGFVLYVWLTCTQQLUESGHLAPGCGPTTAYVYVLSE 187

```

```

DB      318  MEAHENFENRENSQIFVQVQVLDQKMWATC:KEEIEKGRDPYSICKCTTSINTTTFEN 377
QY      188  LMQCKPOAFIQGFSPKRGCTGAAAGLWKNQKILFEEQDLRKGLGEGVSFAFNXNIF 246
DE      178  LFTQNAHSEKSKQDLOHLSLAAGMWTFYFGEELRRKSGIMDSQIRHLLDQVLE 437
QY      247  QKDINCEHSEFAL----- 260
DB      416  INHRENSYTFHNVGVEVAAFYLLKSNIVHSPSOEVKSIEKIMFAFKKVVQVQIF 497
QY      261  SRFLPRLNERTSHLEKSLQWYSFHKMELQWISKQKQSDSTLQQSGLEFFSC 320
DE      418  GSFELGLHSEKQFLDAPFHQGLSCIEKRSQVCLLETISQNESLQEGIDCKKILFYCLFE 557
QY      331  IQEEEPQALSHFCVIVVSNIAK--MERVSSFCRKCESAQVHLHYGATYSADGEDR 378
DE      558  NDDDTFAEANKVECI---NFAKCVSDVVAAGCKKHCFTLKLKLSHSTQNVLSGAQEH 614
QY      379  A---KCSAGATILY-----GLPERTVLLDAYS----- 403
DB      615  SYVEFLDAMNHELGSVFTSKDQELRMKDTNLSASFVLYNKLKYNNYTNVLYVANV 674
QY      404  ---HRLAAALCNPNLHLSVYNALSGRVKILCOGLRHPKILQNL 445
DE      675  PVEVEXHFFRLQVWKNQCHLMSITLSISQVLLCDVLSQASQNTIEEL 124

RESULT 5
Q8IXT0
ID Q8IXT0 PRELIMINARY: 187; 546 AA.
AC Q8IXT0
DE 01-MAR-2003 (TXEMBLrel. 23, Created)
DT 01-MAR-2003 (TXEMBLrel. 23, Last sequence update)
DI 01-MAR-2003 (TXEMBLrel. 23, Last annotation update)
DE Similar to NALP2 protein.
CS Homo sapiens (human).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.
CX NCBI_TaxID:9606
KW [1]
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg P.
RL Submitted (May 2002) to the EMBL/GenBank/CCDS databases.
TR EMBL; AF040004; AB012661;
SQ SEQUENCE 546 AA; 61640 MW; 4F8E545AC319F7 CARY4.

Query Match 23.7%; Score 549; DP 5; Length 546.
Best Local Similarity 30.1%; Pred. No. 190 44;
Matches 151; Conservative 145; Mismatches 181; Indels 85; Gaps 13.

QY      8  CSVATGCTSCWPSPAPQLQELIRVFERLITIDGFDELKSFQCPQPKLQWEEKRPTE 67
DE      256  CSPAEHNRWPKPELQCDRPHILACAKKILFVLDGFDLGAAPGALIEDICQENKRRKVP 215
QY      68  LKNSLIRKLLPELSLITITRTALEKRLHLSHPHREVLTPSPAERKRYFYFYFNA 127
DE      316  VLGSLIRKVVLPKALIVITRPEADPCRLTAEPIVIRVEGFLEEDRAYFLRHGDE 375
QY      128  EQACQVENVYRCNEPFLTMGFVLYVWLTCTQQLUESGGLLRQTSRTTAVVYVLSE 187
DE      316  EQAEAFELKPSNAALFQGSAPAVCWIVTTLKQWEGEDPVPTCLTRTGLFELFSCS 425
QY      188  LMQCKPOAPRQPPNCRKGLSIANKGLKILKLEBQDLRKHLGDEGVSFAFNLMNIFQ 247
DE      416  PF---EQCAQIRGA--LRTISLAAQGLWAQTSVLRHEDLERLQVQESDLRLFLDGLIR 490
QY      248  KD-IMPERSFLADT-SRPT----- 264
DE      491  QCPVQAGTYSFELHSGCCQFTALFYTLKKEEELRQCHTMCIDGVCKLLSGVERLNPD 549
QY      259  ---FRLNERTSHLEKSLQWYSFHKMELQWISKQKQSDSTLQQSGLEFFSC 314

```

```

DB 500 LQAGYVSYGLAVEKRAKELEATEFEMVDFPDKORLEEDINCRGHSVTVDQ 605
QY 315 FSLVLEQEEFTQQALSIFQVWNSINASKMEHYVSGFELZKCSAQWHLV 667
DB 606 LQGLYSOEELKEVYACFKK SLENAVAVPSSSTVAPRPAIQKYSQVAFENLP 663
QY 469 GATYSADHEDPARCSAGANTLAVCLPAPKTVLDAYSERHAAALCTNPKL 426
DB 664 ENVTASEDAEVSQDQCH KLPPFWTGLCSFG SSKDGLAINESEF 711
QY 427 GSRGVKGLGGLRHPNKLNL 448
DB 712 SASLVLCEQIASOTCHLGRV 733

RESULT 6
Q99XWC PRELIMINARY; PRT: 561 AA.
AC Q99XWC;
DT 01-MAR-2003 (Tremblere), 23, Created;
DT 01-MAR-2003 (Tremblere), 23, Last sequence update;
DT 01-MAR-2003 (Tremblere), 23, Last annotation update;
DE Hypothetical RN1-like structure containing protein;
DS Mus musculus (Mouse);
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
OX NCBI TaxID:10090;
RN 1;
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=OVARY;
RX MEDLINE=22354681; PubMed=12446481.
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase 1 & 2 Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
PT 62,770 full-length cDNAs";
RL Nature 422(693):573(2002);
DR EXBL: AK054426; SACS5775;
KW Hypothetical protein;
SQ REFERENCE 863 AA; 93955 MW; 614649820544 P 1047;

Query Match 21.8%; Score 519; DB 11; Length 863;
Best local similarity 24.6%; Pct. No. 100-41;
Matches 149; Conservative 91; Mismatches 160; Indels 180; Gaps 10;

QY 9 SSKLIFSCWFEESAPQCELLVIRERLPTTQSGFIFTHHFGJPMFQWCEKPTFL 69
DB 194 SLAMTSRWSPSAPWEEISQEKILPTDLEAEWAVLQQAQVWCEKGVAV 253
QY 60 LLSLIRKKLLPELSLITNTTALERHGLSHHGHVPEHLSHAKRKYFYKYNHAR 128
DB 254 LLSLIRKKLLPESSLLISTSSETERELKQWELVTHQPTTHKKNINMCFHSLEQDN 313
QY 129 QACQVNVVRCNPLEFTMGVPLVQWVWVTVLQGLGGLLPLSRTTAVVLYLSL 188
DB 114 LAEAFSLRNECLFTVCAPVWVYVATLKHESLRHNSISERTTSYTTILNL 373
QY 189 MQPKSAPRLQPPNCRGLASLAADKMSKRIKLEELPLPKFGLGDSVAFIRM 243
DB 374 FIFR-- NAQNSNSEDCLNLPFLAVEMWATDQVGENRALPANGWDSITLDDI 429
QY 244 NIFQKQINCERSFAL-- 259
DB 430 GLEQSESENSYFLRPSVQCFVNAWPLYLHGHETHSNQVYFETFLFLNKKQW 489
QY 260 --TSSEFLGLNEETRSHLEKSLFWKYSRHRKMLLQNLQKAKSGDSTQCGSL 316
DB 490 VFLOCFEFLHETEQEKLEAPFTHLKRFRGDFPWLSEILITLHFWKNTWKFFY 549
QY 117 CVYKIQEEFTQQA-----LSHFTVIVVSNIAK----- 345
DB 550 CLPMESEEVFQSAVNCHEQILVVVKDVSQVFNVAWVLSQAGALTFSTSAQV 605
QY 346 -----MEHVS 351

```

```

DB 610 QPGKIMLWHQICSVF.PNKQIKTLRIEDTTFNEPVFKIFYVLEKSSCLKTLVAVNS 669
QY 352 SFCLKA-----CRSAQVHLHYGATVSADGE-----DKAKSAGAH 387
DB 410 PCKCKSLSEHLIQSHYHFFILVLTFLSHSVFM.CGLNQAE"NIALLDLANKSLCHS 729
QY 388 LVAVLPETVLLDAYSERHAAALCTNPKLFEISYRNALSGRVKLLCOOLRHPNKLQN 447
DB 430 W-----VYLSVLPONKSDRYNLISYNNKDKESKALCPALTLPSALHS 774
QY 448 1 448
DB 775 1 775

RESULT 7
Q99XWC PRELIMINARY; PRT: 748 AA.
AC Q99XWC;
DT 01-JUN-2001 (Tremblere), 19, Created;
DT 01-JUN-2001 (Tremblere), 19, Last sequence update;
DT 01-MAR-2003 (Tremblere), 23, Last annotation update;
DE Ribonuclease (angiotensin inhibitor);
DS RNP2;
CC Mus musculus (Mouse);
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
OX NCBI TaxID:10090;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=21757048; PubMed=12273525;
RX Wang P.J., McFarley J.R., Yang F., Page D.C.;
RT "An abundance of X-linked genes expressed in spermatogenesis";
RL Nat. Genet. 23:422-426(2001);
DR EXBL: AF255581; SAKS1961;
DR HSSP: P10775; 2BRH;
DR XGR: NG011896018; RHL2;
DR INTERPLOT: IPR007091; LRR_PNIN;
DR PROSITE: PS50503; LRR_R1;
SQ SEQUENCE 748 AA; 84906 MW; 6741445E01D58E8 CRC64;

Query Match 20.8%; Score 494; DB 11; Length 748;
Best local similarity 26.3%; Pct. No. 4.2e-39;
Matches 109; Conservative 90; Mismatches 171; Indels 94; Gaps 9;

QY 77 LFWKEEETELHLSNLSLRKKLLELGLTTRTATCEKLHLEHREVELIGPSEAF 116
DB 13 VDTATETVPLILANSLRRKMLKSGFHSATPTTFQMEGRVCTNVKIVTFENESI 72
QY 117 KEVFKYFENARQAGVENVYVDFNEFLPTNGFVPLVQWVWVTCQGLGSGGLRQTSRT 176
DB 43 KMYFSLFQERTKTQFHSANKENQOFLTVQVPLQWVATLCKKEIKRGDLVSVCR 132
QY 177 TTAVVLYLSLQMPQKQAFLQPPNCRQRCGLSLAUGLWKNKILPBEQOLRHKGLDGE 235
DB 113 TSLATTHPLPLFPQSAQYPSKESQALQSCSLAASQVMTDTFVFGEEALRRNGMD 192
QY 234 DVSATLKNVIFCKQINKESSEFAL----- 259
DB 193 DPTPLDVLLEKSKKSKSYFLHPSIQEVCAIFYLLKSHMHPSCVYSIALITFT 252
QY 260 -----SRPFGILNEETRSHLEKSLKWKVSHIKMDLQWIOSKAGSDSTLQQ 309
DB 253 LKKVAVGVLFQSGPFLHLESEQKLEAFQCHQISQIKSQLYQCLFTISNSELQFQV 312
QY 310 LMEFPLLYLQCEEFQCALSHFQVTVSNIAKMEHVSFCLKPRSAQVHLV 367
DB 413 DQKILFYLFEMCDERFLAQAKNMEQIFPAKQVSDVIVAAHCLQHSTLKKSL 369
QY 568 IATYDRIEDVA-----RGS--AGAHLLVQLP-----ERTVLLDAYSERHAAA 409

```


Best Local Similarity 22.28; Pref. No. 1.1e 10;
Matches 124; Conservative 39; Mismatches 210; Indels 146; Gaps 27;
CY 9 SMOGLIES--CWPE PSAPQOEIVFVHRLLEFDEHLEKSPHDPQPKCL7 WEEK 63
DB 2 SIOQLLEHFCYPPDEEVESFLRTHCALFTFDEHLEKSPHDPQPKCL7 WEEK 63
CY 64 RPTLELINSCKPKLSPLESLITRTALEKHLHLEHPRHVEILQPSRAEKKEFYKY 123
DB 62 HULVLJAN-LUSRLKAGKLLTAR TQVEYPPQLLR KKVLKQSPSHSLRAYARR 117
CY 124 FHNACAGQGVVYVNEHPRFVMEVPLVWVYVCTCLOQ---QLES-GGLLRQTSPTTA 179
DB 118 FFEETACHELQQICAKPNLSLSCVPLVWVYVCTCLOQ---QLES-GGLLRQTSPTTA 179
CY 180 VYMLYL-LSLYCPKPGA--PALQTPNQ-----RSLAAGLKKOKL7 WEEK 63
DB 178 VFLVTVLHNRPOSSLVQNTSPATLRAKPTLHAGLGVHAGTCKSLVFGQEV 246
CY 228 RKHGLDGE-----VSAPLN-----MNIF 246
DB 218 QASKLOGLDGLPALPOVGHVQSYVHPRHLCQAFVTAFLVADKYSTRELLRFF 297
CY 247 QK-----DINCERSHIA-----VSAPLN-----MNIF 246
DB 298 RENTSPREATSSSSCHSSPSPQATLQPSRIRHPRHPRHPRHPRHPRHPRHPRH 356
CY 277 EKSLCK-----VSHIKYELLOWQCS--KAGSDGSTLQSGSDEF-ESLYEIQ 322
DB 357 DRQLVPAKILRRKKA7WAGLFASIRYKSLVAVGUSNCHVNTFFMMHLCITTC 416
CY 313 FEEPIQALSHFOVIVVSNIAK--NEHVYSEPTLERMS--AQVHLVQATYSAD 374
DB 412 SKK-----VSLAARSLAGVLEKACVAGNAGSALSAPV-HHPRQLALD 769
CY 478 TEGARSASANTLQACLPETVL--CAQVHAAAL--TNPCLIE LSLYRNALG 427
DB 463 LNNELNRYVQVQIQPSKATVIVLQNTATLQVTVTEELFYKIVITFLDYNNQIT 522
CY 424 SRVKKL-----QQLRH 440
DB 523 QIARYVAQLDGRJGRH 541

RESULT 15
CASHB2 PRELIMINARY, F50
AC CASHB2
CT C1 MAR 2003 (Tremborel, 23), Tremborel
IT C1 MAR 2003 (Tremborel, 23), Tremborel
NT C1 MAR 2003 (Tremborel, 23), Tremborel
DE Hypothetical protein C130075P1, similar to GAD65, a putative domain
UE protein 47.
CS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Placental; Rodentia; Muridae; Murinae; Mus.
CC Mammalia; Euthalia; Rodentia; Sciuridae; Muridae; Murinae; Mus.
CX MBL:tax10-10090;
RN
RP SEQUENCE FROM N.A.
RC Strains: FVB/N;
PA Strausberg P.
RN Submitted (JAN 2003) to the EMBL/GenBank/DDBJ databases.
R2 SEQUENCE FROM N.A.
RC STRAIN: C130075P1; Spleen;
PX NCBI:1572345683; PubMed:11464951.
PA The EMBL Consortium,
RT Analysis of the mouse transcript (time taken on terminal) annotation
PT 40,770 full-length cDNAs.
PI Nature 420:563-573 (2002).
LR EMBL: AC042670; AAC42671;
RF EMBL: AC042670; AAC42671;

KW Hypothetical protein.
SQ SEQUENCE 553 AA; 107739 MW; 390639621CEB1A58 CRC64;
Query Match 9.84; Score 233; DB 11; Length 953;
Best Local Similarity 22.28; Pref. No. 1.1e 13;
Matches 124; Conservative 39; Mismatches 210; Indels 146; Gaps 27;
CY 9 SMOGLIES CWPE PSAPQOEIVFVHRLLEFDEHLEKSPHDPQPKCL7 WEEK 63
DB 243 SIOQLLEHFCYPPDEEVESFLRTHCALFTFDEHLEKSPHDPQPKCL7 WEEK 63
CY 64 RPTLELINSCKPKLSPLESLITRTALEKHLHLEHPRHVEILQPSRAEKKEFYKY 123
DB 62 HULVLJAN-LUSRLKAGKLLTAR TQVEYPPQLLR KKVLKQSPSHSLRAYARR 164
CY 124 FHNACAGQGVVYVNEHPRFVMEVPLVWVYVCTCLOQ---QLES-GGLLRQTSPTTA 179
DB 118 FFEETACHELQQICAKPNLSLSCVPLVWVYVCTCLOQ---QLES-GGLLRQTSPTTA 179
CY 180 VYMLYL-LSLYCPKPGA--PALQTPNQ-----RSLAAGLKKOKL7 WEEK 63
DB 178 VFLVTVLHNRPOSSLVQNTSPATLRAKPTLHAGLGVHAGTCKSLVFGQEV 246
CY 228 RKHGLDGE-----VSAPLN-----MNIF 246
DB 298 RENTSPREATSSSSCHSSPSPQATLQPSRIRHPRHPRHPRHPRHPRHPRH 356
CY 277 EKSLCK-----VSHIKYELLOWQCS--KAGSDGSTLQSGSDEF-ESLYEIQ 322
DB 357 DRQLVPAKILRRKKA7WAGLFASIRYKSLVAVGUSNCHVNTFFMMHLCITTC 416
CY 313 FEEPIQALSHFOVIVVSNIAK--NEHVYSEPTLERMS--AQVHLVQATYSAD 374
DB 412 SKK-----VSLAARSLAGVLEKACVAGNAGSALSAPV-HHPRQLALD 769
CY 478 TEGARSASANTLQACLPETVL--CAQVHAAAL--TNPCLIE LSLYRNALG 427
DB 463 LNNELNRYVQVQIQPSKATVIVLQNTATLQVTVTEELFYKIVITFLDYNNQIT 522
CY 424 SRVKKL-----QQLRH 440
DB 523 QIARYVAQLDGRJGRH 541

Search completed: Sat Oct 31, 2003, 10:07:46
Job time: 114 secs